

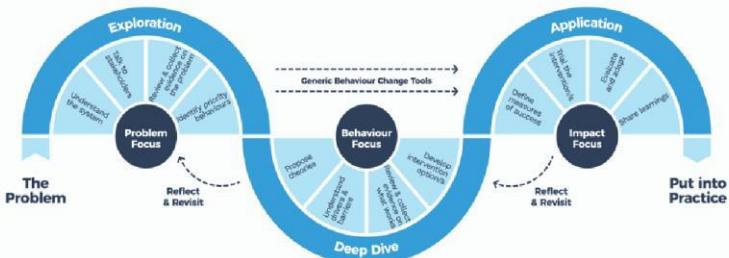


THE METHOD

CHAPTER 1: Evidence Reviews *Connecting Research Knowledge to Practice*

A/Prof Peter Bragge, Alyse Lennox, Dr Jim Curtis,
Mark Boulet, Dr Peter Slattery, Dr Julia Meis-Harris, Abby Wild

www.behaviourworksaustralia.org



INTRODUCTION

How do we stop people from calling an Ambulance when they don't really need one? This was the question faced by the Victorian Department of Health and Human Services in 2016. Calls to Ambulance Victoria were rising faster than population growth. Over 40,000 of these calls - representing almost one in 10 – ended up not requiring an emergency ambulance response. The resource cost of triaging and managing these calls was threatening timely delivery of services for genuine medical emergencies. A previous campaign aimed at a key demographic who were found to be frequent ‘non-emergency’ ambulance callers had failed.

Understanding and influencing human behaviour has been the focus of decades of applied psychology research and centuries of philosophy and other forms of scientific enquiry. This means that challenges like reducing unnecessary emergency calls have been encountered and studied the world over. Knowledge from this research may not give ‘the answer’ - but it can add considerable value to problem-solving by illuminating potential solutions - and just as importantly, flagging areas where investment may be wasted on strategies that have failed elsewhere.

This chapter outlines how evidence reviews can assemble this knowledge and make it useful to policymakers and other leaders who need to make decisions and take action to solve problems. Before diving into review processes, it's important to clarify some key definitions.

Any investigation that produces new knowledge, understanding or insights can be considered “research” (The National Health and Medical Research Council, Australian Research Council, and Universities Australia 2018). This includes activities described in other ways such as ‘data insights’ or ‘business intelligence.’ However, University research that is peer-reviewed, published in academic journals, and subject to ethical standards is widely considered to be the most rigorous form of defensible knowledge available. Accordingly, we will consider universitybased, peerreviewed research to be evidence in this chapter. It is important to acknowledge that other sectors can have a much broader definition of the term ‘evidence’. For example, governments and the media may refer to evidence as testimony of an expert, findings of an audit, information from a consulting report or anecdotal evidence gathered from people around them. These sources are not explored within the content of this chapter.

Grey literature is the term given to documents not controlled by academic publishing organisations and therefore published outside of peer-reviewed journals (Adams et al. 2016). Government and other organisational reports, working papers, ‘white’ papers and evaluations are all examples of grey literature. These publications can contain valuable information on innovations and interventions that are developed, tested and evaluated outside of university settings. Grey literature is not peer-reviewed in the same way as published academic studies and is therefore considered less robust (acknowledging that the university research model of peer-review is not perfect).

WHAT IS AN EVIDENCE REVIEW?

An evidence review is any effort that brings together multiple individual research studies in a defined area in order to consolidate knowledge across these studies. In doing so, evidence reviews paint a more reliable picture of knowledge. For example, a single research study may report that drug A is more effective than drug B. However, nine other studies may find the opposite. An evidence review brings all of these together to give an overall picture of all studies of drug A versus drug B. By doing an evidence review and analysing all the studies, it may be that the one 'outlier' study is poorly designed, has a very small sample size, or is conducted in a different setting to the other nine studies. If a pharmaceutical regulator or government relied only on that one study, they may waste investment in drug A. An evidence review of all 'drug A versus drug B' studies prevents this error. For this reason, evidence reviews are the preferred unit of knowledge translation into policy and practice (Grimshaw et al. 2012).

There are three key types of evidence review that are the focus of this chapter - systematic reviews, rapid reviews and narrative reviews.

Systematic reviews

Systematic reviews are an overview of individual, or primary studies "which contains an explicit statement of objectives, materials, and methods and has been conducted according to an explicit and reproducible methodology" (Greenhalgh 1997). Systematic reviews are therefore known as secondary studies. Systematic reviewing has been around for hundreds of years, but the modern systematic review grew out of the establishment in the early 1980s of the Cochrane Collaboration, established by Archie Cochrane to focus on creating and disseminating reviews of all known studies on specific healthcare interventions (Shah and Chung 2009).

Undertaking a robust systematic review of all known individual studies in a particular topic area can take many months or even over one year. Systematic reviews are therefore well suited to in-depth explorations of relatively narrow topic areas, such as 'the effectiveness of drug A vs. drug B for treating condition X.' They are often used as an entry point into PhDs and other university-based research, where a deep understanding of the specifics of individual studies and their shortcomings is required to build knowledge by identifying and filling research gaps.

Rapid reviews

Rapid reviews use the same approach as systematic reviews but instead of bringing together *primary* studies, rapid reviews bring together *secondary* studies - that is, they are 'reviews of reviews.' Rapid reviews evolved in the 21st century in response to increasing demands for research inputs into policymaking in shorter time frames than systematic reviews. Due to the pressure of electoral cycles and sometimes limited public service resources, policymakers need answers in days or weeks; not months. Furthermore, the systematic review 'grunt-work' has already been done, so rapid reviews can pull together the headline findings in less time (Khangura et al. 2012). There is some evidence to show that despite the differences in approach between rapid and systematic reviews, their essential conclusions are similar (Watt et al. 2008).

Rapid reviews have served policymakers well in addressing pressing questions in (almost) real time and are the dominant review approach used as part of the BehaviourWorks Method. They generally address broader questions than systematic reviews which also aids their application to policy making. However, systematic reviews remain the most comprehensive approach to gathering in-depth understanding of research in a specific domain, and should always be considered where resources and timelines allow.

Narrative reviews

Like systematic and rapid reviews, narrative reviews bring together a body of research evidence pertaining to a specific topic or topic area. However, the methods by which research studies were searched for and selected are not made explicit, and narrative reviews do not evaluate included studies for methodological quality (West et al. 2002). This means they are less structured and more prone to bias - for example they may select research that supports a particular perspective or point of view and omit other research. Therefore, narrative reviews should be interpreted with caution. The easiest way to tell the difference between a narrative review is to look in the methods section - if there is no “*explicit statement of objectives, materials, and methods*” – especially how the review looked for and selected studies - it is probably a narrative review.

As this chapter focuses on systematic approaches to reviewing, the remainder will focus on these types of reviews.

WHAT IS INVOLVED IN DOING A REVIEW?

All systematic and rapid reviews involve a **core set of components** and associated activities based on well-established, universal review methods. Some reviews also contain **additional components** which address specific needs and add value in various ways. Table 1 describes the core components (shaded rows) and additional components (non-shaded rows).

Table 1: Components of systematic and rapid reviews

Component	Description
Background	Outlines the topic, key definitions and reason for the review
Review question	The specific question being addressed by the review

Review scope: inclusion and exclusion criteria	Specifies what is in and out of the review based on the characteristics of the desired studies (e.g. study type, design, focus) and other factors (e.g. year published, English language vs. other)
Review sources: search strategy	Describes where the review information will be sourced (e.g. academic databases, grey literature, websites and web-based platforms)
Key information from included studies	Presents (usually in a table) the key characteristics and findings of the included studies
Quality of included studies	
Meta-analysis	Provides an indication of how robust the included studies in the review are based on their adherence to accepted research methods and processes. Enables the level of confidence in the review findings to be established
Non-report outputs	A mathematical estimate of the overall effect of an intervention across multiple studies. Enables the overall effect to be quantified, taking into account variability across individual studies. Meta-analyses are beyond the scope of this chapter and is extensively described elsewhere (Deeks, Higgins, and Altman 2019) Additional representations of the review findings (e.g. evidence map / grid, network map, infographic) that can enhance engagement with the review audience and therefore enhance impact.
Practice review	One-on-one interviews with leaders and practitioners in the field to ascertain 'what's happening on the ground' in the local context. Enables real-world perspectives on practical application of review findings by exploring acceptability, feasibility, transferability, cost-effectiveness, scalability and other information generally not found in published studies.
Summary of findings	Presents the key findings of the review based on the dominant characteristics and themes across the included studies
Commentary	Description of the implications of the review findings for policy and / or practice based on the context of the review. Adds an interpretation of the meaning of the review findings that is tailored to the review client. This may include recommendations for future research but generally stops short of practice recommendations, which do not fall within the responsibility or authority of the review team
Reference list	Complete list of all references including references to all included studies
List of related studies	Studies are often found that do not directly address the review question but may be of interest. These can be presented as an additional reference list, which may include brief annotation or a verbatim copy of the studies' abstracts

Methods	Description of the review approach and procedures (generally an appendix if the review is not produced for academic publication)
---------	--

THE REVIEW NAVIGATOR

Ensuring that reviews are appropriately tailored to the needs of end-users (*clients*) and matching their time and resource budget involves a number of decisions. In the course of producing hundreds of reviews since 2011, we have learned that given the resource-intensity of the review process, failing to consider these decisions early can result in considerable wasted effort.

Therefore, the remainder of this chapter presents a *review navigator* based on principles that are critical to *any* review - from a 1-day overview of reviews right through to a large systematic review that takes over a year. The navigator presents a series of questions that are grouped under these principles. Some of these are contained in or based upon the STARR (SelecTing Approaches for Rapid Reviews) decision tool, a freely available guide to review planning (Pandor et al. 2019; The University of Sheffield 2020). Not all of the questions are necessarily relevant to all reviews - some information may already be known. The answers to the questions will form the methodological backbone of the review, based on specific needs, timelines and resources available. We'll expand on the case study presented in the introduction to illustrate the review navigation process.

PRINCIPLE 1: Understand the context and scope

Understanding the context and scope of a review is pivotal to establishing every subsequent step of the process. The following questions are designed for a review team and client to reach a shared understanding of the review rationale, audience and intended use. The questions may prompt the need to gather further information before proceeding.

Who is the review for?

This gives critical information about the likely influence of the review findings. In an organisational setting, seniority and decision-making authority of the review audience is broadly proportional to the review's level of importance and priority.

What does the review client plan to do with the review findings?

The intended use of the review signals its level of potential influence. For example, a background review to inform an internal organisational strategy has less immediate impact than a review to inform an urgent policy decision such as whether to enforce mask-wearing to control COVID-19.

Are there any relevant background materials?

It's very helpful to access background papers, reports, previous reviews or other information developed or known to the client. This can provide valuable information about topic definition, context and rationale - and if a similar review has been commissioned or undertaken in the

past it could save considerable time and effort in building background definitions and developing the review method.

Aside from this review, what are the other known decision-making inputs?

Reviews are one of many inputs into decision-making. These include (but are not limited to) data, statistics or other information pertaining to feasibility, acceptability, ease of implementation, timeliness, scalability, cost / return on investment and (in policy contexts) political climate relating to interventions or decisions (Lavis et al. 2004). Each of these inputs is weighted differently both by individuals and also across problem areas of various size, scope and desired impact. Knowing about these other inputs is helpful in framing the context for the review and the decisions or processes it is feeding into.

When does the review need to be completed?

If action or a decision is required based on the review by a certain date, this directly influences resource allocation and approach. For example, timeframes under 4 - 6 months generally suggest a *rapid review* is the only feasible option, while very short timeframes (less than a week) usually mean that a practice review is not possible.

Will it be a rapid review of reviews, or a systematic review of primary studies?

As discussed above:

- Systematic reviews are best suited to in-depth explorations of relatively narrow topic areas and take at least 3 - 4 months
- Rapid reviews generally address broader questions; are useful for gathering headline findings to inform policy or other decision-making processes; and can be done in short (weeks) or very short (one day or less than a week) timeframes

What level of confidence in review findings is required?

Critical decisions and / or large financial investments need to be justified and, at times, defended. There are a number of techniques that can maximise the robustness of a review and therefore, the level of confidence in the review findings. These include:

- Using two reviewers to independently evaluate what is 'in' and 'out' according to the review scope - this reduces bias in study selection;
- Undertaking a formal quality appraisal of the studies included in the review and weighting the review findings and interpretation based on this; and
- Expanding the search parameters (for example, searching more databases or across a longer time span) to minimise the chances of not identifying relevant studies.

What is the desired report format?

Reviews are usually presented in the form of a written (and sometimes long) report. Where a report is desired, it is helpful to know the preferred format (for example a 1-page executive summary, 3 pages of key findings in dot point form, main report text, appendices). This sets the minimum output and also reduces wasted effort on long reports or tables that may not be useful to the client. Although written reports are the default output, there are a range of

non-report outputs including a slide deck or visualisations such as maps / grids, infographics and even animations. If these are requested, the additional costs (for example, use of graphic design or information technology expertise) need to be considered.

A key and often-neglected issue is ensuring there is a shared understanding between the reviewers and the audience. Use of plain language rather than academic or specialist terminology, circulation of a draft report prior and an interactive presentation with a question-and-answer session are all useful strategies for ensuring that the review maximises its impact. It is important to flag this early and build time in to accommodate these key activities. A written report handed to a client without this discourse carries risk for the review team (findings being misrepresented and / or misused) and the client (taking action based on a misunderstanding of key information). This is further explored [here](#).

Based on the above questions, what level of resourcing is needed?

The above questions should generate information on the people and tasks involved in the review and enable an estimate of person-days required to complete the review within the specified timeframe. If this resource cost exceeds the available budget, discussion of the cost / benefit of various approaches and alterations in light of this will be required.

Worked example: Context and scope

Here's how the above questions apply to the example presented earlier of the challenge: "*How do we stop people from calling an Ambulance when they don't really need one?*"

Question	Application to ambulance review
Who is the review for?	<ul style="list-style-type: none">The review was for the Victorian Department of Health and Human Services (DHHS)
What does the review client plan to do with the review findings?	<ul style="list-style-type: none">The review was designed to inform a strategy to address the problem of rising numbers of inappropriate calls to Victoria's Ambulance Service
Are there any relevant background materials?	<ul style="list-style-type: none">Victoria's Ambulance Action plan provided data on call volumes to the ambulance service in Victoria (State of Victoria 2015)Previous research commissioned by DHHS had examined the drivers of inappropriate ambulance use, providing information about the nature of the problem. This included the development of typologies - groups within the community defined by their attitudes to when it is appropriate to use an ambulance

Aside from this review, what are the other known decisionmaking inputs?	<ul style="list-style-type: none"> The review was designed to inform a structured stakeholder dialogue with senior representatives from the Ambulance Service, the Victorian Government, the Emergency Services Telephone Authority and other key stakeholder groups
When does the review need to be complete?	<ul style="list-style-type: none"> The review had a timeline of approximately one month
Will it be a rapid review of reviews, or a systematic review of primary studies?	<ul style="list-style-type: none"> A rapid review approach was employed due to the short timelines and review context
What level of confidence in review findings is required?	<ul style="list-style-type: none"> A high level of confidence in review findings was required, as the review was informing a major statewide government initiative For this reason, a comprehensive database search and quality appraisal of included studies was required
What is the desired report format?	<ul style="list-style-type: none"> A written report was required, including questions for deliberation in a subsequent day-long, multi-stakeholder dialogue (stakeholder dialogues will feature in a later chapter of the BWA Method Book)
Based on the above questions, what level of resourcing is needed?	<ul style="list-style-type: none"> Given the need for quality appraisal and interviews (including an ethics application), three researchers were involved in the review

PRINCIPLE 2: Get the research question right

“Judge of a man by his questions rather than by his answers” (Voltaire, French author, humanist, rationalist, & satirist; 1694–1778). (“Quotations Page by Author” 2020)

It is imperative to fully understand the question being addressed by any review for two reasons. First, this prevents unnecessary use of resources in the event that the question needs to be reframed at a later date; second, it is often the case that clients will ask for reviews that encompass multiple questions. Given the work that goes into developing protocols for a single question, it is often necessary to prioritise the most important question, then return to second- and third-order issues as resources allow.

What is the primary decision question for the rapid review?

Often, review clients want to know what works (for example, ‘does pay-for-performance work to improve healthcare quality?’), or decide between various intervention approaches (‘which is more effective - pay for performance or implementation of best-practice guidelines?’).

While “what works?” is an understandably dominant question, there are many others. Table 2 presents examples of these questions; illustrates the type of research that addresses these questions; and shows how the example of “*How do we stop people from calling an Ambulance when they don’t really need one?*” can be altered according to the focus of each question.

Table 2: Getting the question right

Question	Type of research that addresses the question	Examples of how this question would apply to the ambulance review
What is the problem?	Observational studies - can be numerical (e.g. how many people get this disease, are rates going up or down) or non-numerical / qualitative (e.g. what does it feel like to go through cancer treatment, divorce, losing a job)	How many Victorians dial an ambulance when they don't really need one? What are the demographic and other characteristics of people who frequently dial and ambulance when they don't really need one?
Does this intervention work?	Effectiveness of various interventions (e.g. employment coaching, management of chronic health conditions, drug rehabilitation) aimed at reducing homelessness How - key features of successful approaches, barriers and facilitators of success	What are effective strategies to reduce inappropriate ambulance calls? <i>This was the question used for the review</i>
How might this work?	Theories of change based on ‘first principles’ (not real-world data)	What behavioural psychology techniques could reduce the behaviour of calling an ambulance when it is not required?
What risks are present or possible?	Modelling / Forecasting e.g. bushfire risk	If unnecessary ambulance calls keep rising at present rates, what implications does this have for service delivery, costs and adverse outcomes in the future?
What are the standards of practice in this field?	Guidelines based on evidence and input from an expert panel	What guidelines exist for triaging calls to ambulance service?

What approaches are used to ensure that programs, services and products get to those who need them?	Implementation - examining strategies to ensure that interventions are reaching their enduser targets	How are ambulance services accessed and promoted to citizens?
---	--	---

Is the review viable based on this question?

A key step prior to finalising the review plan is determining the viability of the research question. This involves a preliminary examination of the evidence base prior to finalisation of the review question - for example by undertaking a simple preliminary database search or other form of literature scan, examining known research in the area, or consulting experts (The University of Sheffield 2020). In simple terms there are three scenarios that can result:

1. The volume and type of research identified appears appropriate to the review resources and timeframe.
There is no 'magic number' of reviews or primary studies that is 'right' - the review team needs to make a judgement based on a preliminary examination of initial search outputs. For example, the ambulance review identified 5 systematic reviews and 1 narrative review. Had there been no reviews identified, alternative strategies would have been explored.
2. Very little or no relevant research is identified. In this scenario, the review parameters need to be expanded. This could be by expanding the search strategy or the number of data sources, including other study designs, searching across more years or a wider geographical region (if this has been restricted), or expanding the scope to related bodies of literature (for example, looking at emergency services including but not limited to ambulances).
3. The volume of research identified is too large. This requires the opposite of the above – for example, narrowing of review scope by search strategy / databases, topic, years or geographic location. Alternatively, restrictions can be made in other ways - for example by focusing on intervention studies that measure a specific outcome.

Decisions on widening or narrowing the scope involve trade-offs and the specifics of these should be openly discussed with the review client and made explicit in the final report.

Worked example: Getting the question right

Here's how the above questions apply to the example presented earlier of the challenge: "*How do we stop people from calling an Ambulance when they don't really need one?*"

Question	Application to ambulance review
What is the primary decision question for the rapid review?	<ul style="list-style-type: none"> The primary focus was the effectiveness of interventions, and the agreed review question was “What are effective strategies to reduce inappropriate ambulance calls?”
Is the review viable based on this question?	<ul style="list-style-type: none"> Following preliminary examination of the literature and consultation with experts, the scope was confirmed as viable based on the following search parameters: <ul style="list-style-type: none"> Focus on reviews and high quality primary studies examining ambulance service use (police, fire or other emergency services excluded) The main interventions of interest were public health campaigns, behavioural interventions for specific subgroups, with dispatch or triage systems also within scope The key outcome of interest was the number of inappropriate or nonurgent calls to ambulance services or medical emergency phone numbers

PRINCIPLE 3: Look for needles, not haystacks

Anyone who has searched the internet knows the feeling of starting with a few simple search words and ending hours later down an obscure rabbit-hole with no idea how you got there. Browsing the internet is fantastic for resolving a trivia dispute or finding a local restaurant. But it is a highly risky and inefficient place to start an evidence review. The vast majority of the internet is not peer-reviewed research. The quality of the information is therefore unknown and unreliable. It may feel productive, but the numerous links and documents that are generated will most likely be of little use.

The knowledge pyramid (Figure 1) provides context for reviews by illustrating three broad domains of knowledge:

- At the bottom of the pyramid is **the internet**. In December 2020 this was estimated to contain 5.48 billion pages (for regular estimates go to <https://www.worldwidewebsize.com/>);
- In the middle of the pyramid is university-based, peer-reviewed **research articles** with an estimated volume of about 65 million (Jinha 2010; Mabe 2003), or 1.2% of the total pages on the internet. Peer-reviewed research is best found within academic databases which exclusively index articles published in peer-reviewed journals.
- At the top of the pyramid are approximately 200,000 published **reviews**, with around 10,000 being added every year (Clarke and Chalmers 2018), or 0.0036% of the number

of pages on the internet. Reviews can be found in academic databases like Medline, but are also contained in dedicated review databases including:

- Social Systems Evidence, which indexes over 4,000 reviews pertaining to the Sustainable Development Goals (SDGs):
<https://www.socialsystemsevidence.org/about> ;
- Health Systems Evidence, containing over 14,000 reviews examining how health systems are governed, financed and organised to ensure delivery of health services and interventions to those who need them:
<https://www.healthsystemsevidence.org/about>;
- The Cochrane Database of Systematic Reviews, contains over 7,500 reviews of healthcare interventions that are freely available in over 130 countries (Farquhar and Marjoribanks 2019): <https://www.cochranelibrary.com/cdsr/about-cdsr>;
- The Campbell Collaboration, which indexes review-level evidence pertaining to areas including business and management, crime and justice, disability and education:
<https://campbellcollaboration.org/better-evidence.html>; and
- The Collaboration for Environmental Evidence, which focuses on reviews in the areas of sustainability and biodiversity: <https://www.environmentalevidence.org/> .

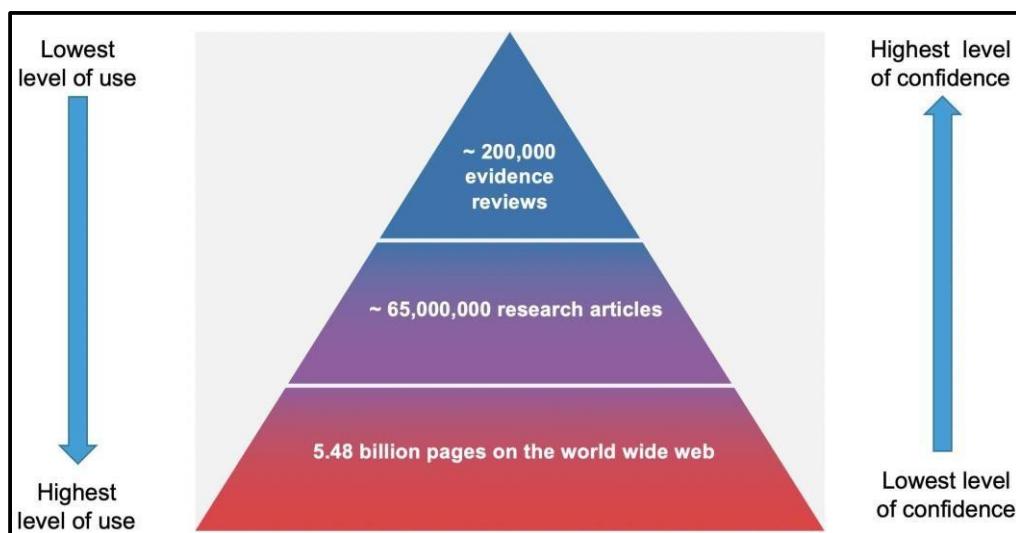


Figure 1: The knowledge pyramid

It seems logical to start looking for research evidence at the very tip of the knowledge pyramid, containing the lowest volume of documents with the highest credibility. Unfortunately and probably due to habit, the opposite seems to be the default. The key to efficient and effective reviewing therefore is to avoid the haystack and focus on the needles. The following questions guide the process of efficient searching for research evidence.

Where will you search for research?

If you are doing a *rapid review* you should target the top of the pyramid - generally by searching a combination of relevant dedicated review databases and academic databases that index

peer-reviewed research. If you are doing a *systematic review* of primary studies you should target the middle of the pyramid, focusing on academic databases and online portals that exclusively list research studies.

How will you search for research?

Although artificial intelligence and other algorithmic and indexing approaches are beginning to simplify searching, it still remains a task that requires input from a specialist literature reviewer or librarian. Most university libraries have freely accessible pages that delve into this technical detail, which is beyond the scope of this chapter. It is strongly advised to seek expert assistance to build a search strategy matching your needs rather than learn the ins and outs of database searching, because even small errors in syntax can ruin an entire search.

Worked example: Look for needles, not haystacks

Here's how the above questions apply to the example presented earlier of the challenge: "*How do we stop people from calling an Ambulance when they don't really need one?*"

Question	Application to ambulance review
Where will you search for research?	<ul style="list-style-type: none">Three academic databases - PubMed, CENTRAL (Cochrane) and Web of Science were searchedThe Google Scholar platform was searched (first 100 results screened)
How will you search for research?	<ul style="list-style-type: none">A search strategy was developed using key terms pertaining to medical emergency telephone numbers; demand for services; dispatch and triage; public campaigns; and inappropriate useEnglish language publications from peer-reviewed journals or grey literature published in the last 5 years were eligible for screening

PRINCIPLE 4: Quality matters

"Put together lots of small rubbish studies and you have a big pile of rubbish studies, not an answer"

(Doctors' Information:

Excessive, Crummy, and Bent" 1997)

A description of studies in a particular area and their key findings can be a useful entry point into understanding knowledge. However, without knowing how well the studies were undertaken, the level of confidence in this knowledge cannot be known. Quality appraisal of the studies relevant to the review question enables this level of confidence to be established. It is highly recommended, especially if the review is informing a major decision or investment as outlined in Principle 1. The following questions apply if quality appraisal is going to be undertaken as part of the review.

How will the quality of included studies be evaluated?

Quality appraisal tools have been developed for most types of research studies, including reviews. Like literature searching, quality appraisal is a specialist skill - however it can be learned, and most quality appraisal tools provide detailed instructions for their use. The AMSTAR tool (<https://amstar.ca/index.php>) is designed for evaluating the quality of systematic and rapid reviews. For primary studies there are a vast range of tools that are tailored to the particular types of research outlined in Table 2. Exploration of these is beyond the scope of this chapter, and it is recommended that a review specialist or librarian be consulted to determine the best tool for specific types of primary studies.

How will the findings of quality appraisal be used to aid interpretation of knowledge?

There are numerous possible outcomes to the quality appraisal process based on the number and type of included studies and the volume of each study type.

In a rapid review of reviews, there are two levels of quality to consider - the quality of the *review* (how well it has searched, selected, appraised and reported) and the *quality of the studies within the review* (as reported by the review). Complicating this further, if the review itself is of poor quality there is less confidence in their conclusions about the research they have evaluated. A simple rule of thumb is to group included reviews in the following four groups, with group 1 being the most value and group 4 being the least:

1. High quality review; included studies reported to be of good quality overall;
2. High quality review; included studies reported to be of poor quality overall;
3. Low quality review; included studies reported to be of good quality overall; 4. Low quality review; included studies reported to be of poor quality overall.

In systematic reviews of primary studies there are three broad possibilities, again in descending order of value:

1. Most of the included studies are of reasonable or good quality
2. There is a mix of good and poor quality studies
3. Most of the included studies are of poor quality

This is a very short introduction to a complex area. A key shortcoming of reviews is that they undertake the quality appraisal but do not go to the next step of using the findings to help understand how reliable and useful the knowledge is. The above guide offers a simplified way of approaching this important task.

Worked example: Quality matters

Here's how the above questions apply to the example presented earlier of the challenge: "*How do we stop people from calling an Ambulance when they don't really need one?*"

Question	Application to ambulance review
How will the quality of included studies be evaluated?	<ul style="list-style-type: none"> The AMSTAR tool was used to evaluate the quality of included systematic reviews
How will the findings of quality appraisal be used to aid interpretation of knowledge?	<ul style="list-style-type: none"> Four of the five included systematic reviews were of high quality and one was low quality This indicated that a moderately high level of confidence could be placed in the systematic review findings A narrative review and eight relevant primary studies were also identified. These were not quality appraised, but information on the findings of these studies was also presented

PRINCIPLE 5: Research doesn't have all the answers

All of the activities outlined so far are sometimes referred to as a 'desktop review' - they can be conducted using a computer with internet access. Desktop reviews have two limitations: first, they can't capture emerging information or trends that are yet to be published; and second, the information is generally not specific to a particular context or setting. For these reasons, a *practice review* can be a useful additional component.

Practice reviews examine 'what's happening on the ground' through one-on-one or small group interviews with people who have experience and / or expertise in the area - for example policymakers, professionals, consumers or representatives of service delivery or support organisations. Their insights - on what has worked and failed in the past, logistical considerations, acceptability of proposed approaches to end-users and what's possible with what resources - enable worldwide knowledge to be filtered through specific problem-solving settings and contexts.

Is a review of practice feasible?

The addition of a practice review adds time and cost to the review process - skilled interviewers are required; time is needed to identify, invite and book participants; ethics approval is required if the review team is university-based; and there are costs associated with interview transcription and equipment. Therefore if the review timelines are under 3 to 4 weeks, a practice review is probably not feasible.

Will a practice review add value?

Assuming a practice review is feasible, consideration should be given to the value that it will add. In most cases, the costs of a practice review are relatively modest in relation to the benefits of adding these perspectives to a desktop review. However there may be exceptions. Examples include on-the-ground issues already being well-known (indeed they may be a driver of the desktop review) and situations where an exploration of practice is occurring as a separate activity (for example, an audit or monitoring and evaluation process). In the latter case, access to findings of any evaluations are very useful in contextualising research knowledge.

Worked example: Research doesn't have all the answers

Here's how the above questions apply to the example presented earlier of the challenge: "*How do we stop people from calling an Ambulance when they don't really need one?*"

Question	Application to ambulance review
Is a review of practice feasible?	<ul style="list-style-type: none"> Yes. There was sufficient time and resources to undertake a practice review, including obtaining of ethics approval from Monash University
Will a practice review add value?	<ul style="list-style-type: none"> Yes. Previous research had characterised the problem of inappropriate calls to the ambulance service, but exploration of possible interventions to address this had not been undertaken Interviews were conducted with five purposively selected individuals with experience and / or expertise in the area of use of Triple Zero emergency ambulance services or large scale public health campaigns The interviews were designed to enable understanding of previous strategies to encourage appropriate ambulance use in Victoria and how they can inform this initiative

CONCLUSION

"Health work teaches us with great rigor that action without knowledge is wasted effort, just as knowledge without action is a wasted resource"

[Lee Jong-Wook (1945-2006), WHO Director General 2003 – 2006] (World Health Organization 2005)

The above quote succinctly demonstrates the value of evidence reviews. They are an opportunity to pause and consider the landscape of knowledge before taking action. Like so many activities in the BWA Method, evidence reviews challenge assumptions and in doing so, often bring great return on investment.

This chapter has focused on five key principles designed to enable those planning reviews to get the most out of a review. *Understanding the context and scope* (Principle 1) is pivotal to

delivering useful information to review clients. *Getting the question right* (Principle 2) forces reviewers and clients to consider the precise question they need to address, and prioritise questions if necessary. The many possible starting points for a review are reflected in the BWA Method, which makes two references to evidence reviews – in the ‘exploration’ phase to understand a problem; and in the ‘deep dive’ phase to ascertain the effectiveness of interventions addressing the problem. *Looking for needles not haystacks* (Principle 3) demonstrates that the most useful knowledge is not found where most people begin their search. *Quality matters* (Principle 4) encourages examination of how well research is conducted; not just what it concludes. *Research doesn’t have all the answers* (Principle 5) shows what even a few structured conversations with people on the ground can add to the written word.

The chapter has focused on principles rather than step-by-step instructions on doing a review, which are readily available elsewhere, including through our own published [blogs](#).

It is important to emphasise that an evidence review can contribute to decisions and actions that may involve substantial investment. It is therefore critical to be explicit about the limits of the methodology employed and the information provided. In addition to the limits inherent in the review process itself, it is also very important to draw a ‘line in the sand’ about the role of reviewers as independent and neutral. The role of a reviewer is to assemble, appraise and present findings of a series of studies in a defined topic area to inform decision-making and action - not to make recommendations as to what the decisions and actions are. Therefore no matter what the final review format is, all evidence reviews should include:

- A section detailing the methodological approach to enable the review methods to be clearly understood and reproduced in any future updates;
- An explicit description of the potential limitations and biases of the chosen rapid review methods. Each of the decisions in a review have implications for the overall strength of the review findings. In addition to discussing the risks and benefits of such decisions, any limitations associated with these should be clearly stated. For example, rapid reviews can assemble research knowledge and make it useful in as little one day, but no practice interviews are possible and, depending on resource availability and the volume of studies, quality appraisal may also not be feasible. Therefore, a 1-day review can *describe* the headline findings of included studies, but cannot make statements about the level of confidence in these findings or how they align with real-world practice.

A final consideration is to encourage clients to stay up to date with emerging research knowledge. Thousands of research studies are published daily and therefore new knowledge relevant to a review will become available almost immediately after it has been completed. Strategies for staying up to date can include email alerts linked to database searches and ‘living’ reviews which are updated regularly by repeating the literature search, selection and summary activities on a regular basis (Elliott et al. 2014).

To read more about the “Ambulances are for emergencies” review and intervention you can freely access the following:

Borg, Kim, Breanna Wright, Liz Sannen, David Dumas, Tony Walker, and Peter Bragge. 2019. "Ambulances Are for Emergencies: Shifting Attitudes through a Research-Informed Behaviour Change Campaign." *Health Research Policy and Systems* 17 (1): 31. <https://doi.org/10.1186/s12961-019-0430-5>

Borg, Kim, David Dumas, Emily Andrew, Karen Smith, Tony Walker, Matthew Haworth, and Peter Bragge. 2020. "Ambulances Are for Emergencies: Shifting Behaviour through a Research-Informed Behaviour Change Campaign." *Health Research Policy and Systems* 18 (1): 9. <https://doi.org/10.1186/s12961-019-0517-z>

Bragge, Peter. 2020. "Four Building Blocks of a Successful Behaviour Change Campaign." *Apolitical*. 2020. https://apolitical.co/en/solution_article/four-building-blocks-of-a-successful-behaviourchangecampaign.

REFERENCES

- Adams, Jean, Frances C. Hillier-Brown, Helen J. Moore, Amelia A. Lake, Vera Araujo-Soares, Martin White, and Carolyn Summerbell. 2016. "Searching and Synthesising 'grey Literature' and 'Grey Information' in Public Health: Critical Reflections on Three Case Studies." *Systematic Reviews* 5 (1): 1–11. <https://doi.org/10.1186/s13643-016-0337-y>.
- Clarke, Mike, and Iain Chalmers. 2018. "Reflections on the History of Systematic Reviews." *BMJ Evidence-Based Medicine* 23 (4): 121–22. <https://doi.org/10.1136/bmjebm-2018110968>.
- Deeks, Jonathan J, Julian PT Higgins, and Douglas G Altman. 2019. "Analysing Data and Undertaking Metaanalyses." In *Cochrane Handbook for Systematic Reviews of Interventions*, 241–84. Wiley. <https://doi.org/10.1002/9781119536604.ch10>.
- "Doctors' Information: Excessive, Crummy, and Bent." 1997. *BMJ* 315.
- Farquhar, Cynthia, and Jane Marjoribanks. 2019. "A Short History of Systematic Reviews." *BJOG: An International Journal of Obstetrics & Gynaecology* 126 (8): 961–62. <https://doi.org/10.1111/1471-0528.15214>.
- Greenhalgh, T. 1997. "Papers That Summarise Other Papers (Systematic Reviews and MetaAnalyses)." *British Medical Journal* 315 (7109): 672 – 675. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9310574.
- Grimshaw, Jeremy M, Martin P Eccles, John N Lavis, Sophie J Hill, and Janet E Squires. 2012. "Knowledge Translation of Research Findings." *Implementation Science* 7 (1): 50. <https://doi.org/10.1186/1748-5908-7-50>.
- Jinha, Arif. 2010. "Article 50 Million: An Estimate of the Number of Scholarly Articles in Existence." *Learned Publishing*. <https://doi.org/10.1087/20100308>.
- Khangura, Sara, Kristin Konnyu, Rob Cushman, Jeremy Grimshaw, and David Moher. 2012.

- "Evidence Summaries: The Evolution of a Rapid Review Approach." *Systematic Reviews* 1 (1): 10. <https://doi.org/10.1186/20464053-1-10>.
- Lavis, John N., Francisco Becerra Posada, Prof Andy Haines, and Eric Osei. 2004. "Use of Research to Inform Public Policymaking." *Lancet* 364 (9445): 1615–21. [https://doi.org/10.1016/S0140-6736\(04\)17317-0](https://doi.org/10.1016/S0140-6736(04)17317-0).
- Mabe, Michael. 2003. "The Growth and Number of Journals." *Serials: The Journal for the Serials Community* 16 (2): 191–97. <https://doi.org/10.1629/16191>.
- Pandor, Abdullah, Eva Kaltenthaler, Marissa Martyn-St James, Ruth Wong, Katy Cooper, Munya Dimairo, Alicia O'Cathain, Fiona Campbell, and Andrew Booth. 2019. "Delphi Consensus Reached to Produce a Decision Tool for Selecting Approaches for Rapid Reviews (STARR)." *Journal of Clinical Epidemiology* 114 (0): 22–29. <https://doi.org/10.1016/j.jclinepi.2019.06.005>.
- "Quotations Page by Author." 2020. 2020. <http://www.quotationspage.com/quotes/Voltaire/>.
- Shah, Hriday M, and Kevin C Chung. 2009. "Archie Cochrane and His Vision for Evidence-Based Medicine." *Plastic and Reconstructive Surgery* 124 (3): 982–88. <https://doi.org/10.1097/PRS.0b013e3181b03928>.
- State of Victoria. 2015. *Victoria's Ambulance Action Plan: Improving Services, Saving Lives*. Melbourne Australia: State of Victoria, Department of Health and Human Services. <https://www2.health.vic.gov.au/hospitals-and-health-services/patient-care/ambulance-and-nept/ambulance-action-plan>.
- The National Health and Medical Research Council, Australian Research Council, and Universities Australia. 2018. *National Statement on Ethical Conduct in Human Research 2007 (Updated 2018)*. Canberra: Commonwealth of Australia. www.nhmrc.gov.au/guidelines/publications/e72%0Ahttps://www.nhmrc.gov.au/aboutus/publications/national-statement-ethical-conduct-human-research-2007-updated-2018.
- The University of Sheffield. 2020. "STARR Decision Tool." 2020. https://figshare.shef.ac.uk/articles/Selecting_Approaches_for_Rapid_Reviews_STARR_Decision_Tool/797_0894.
- Watt, Amber, Alun Cameron, Lana Sturm, Timothy Lathlean, Wendy Babidge, Stephen Blamey, Karen Facey, David Hailey, Inger Norderhaug, and Guy Maddern. 2008. "Rapid versus full systematic reviews: Validity in clinical practice?" *ANZ Journal of Surgery* 78 (11): 1037–40. <https://doi.org/10.1111/j.1445-2197.2008.04730.x>.
- West, S., V. King, T. S. Carey, K. N. Lohr, N. McKoy, S. F. Sutton, and L. Lux. 2002. "Systems to Rate the Strength of Scientific Evidence." *Evidence Report/Technology Assessment (Summary)*. <https://www.ncbi.nlm.nih.gov/books/NBK11930/>.
- World Health Organization. 2005. "Bridging the 'Know-Do' Gap Meeting on Knowledge Translation in Global Health." Retrieved September 25.

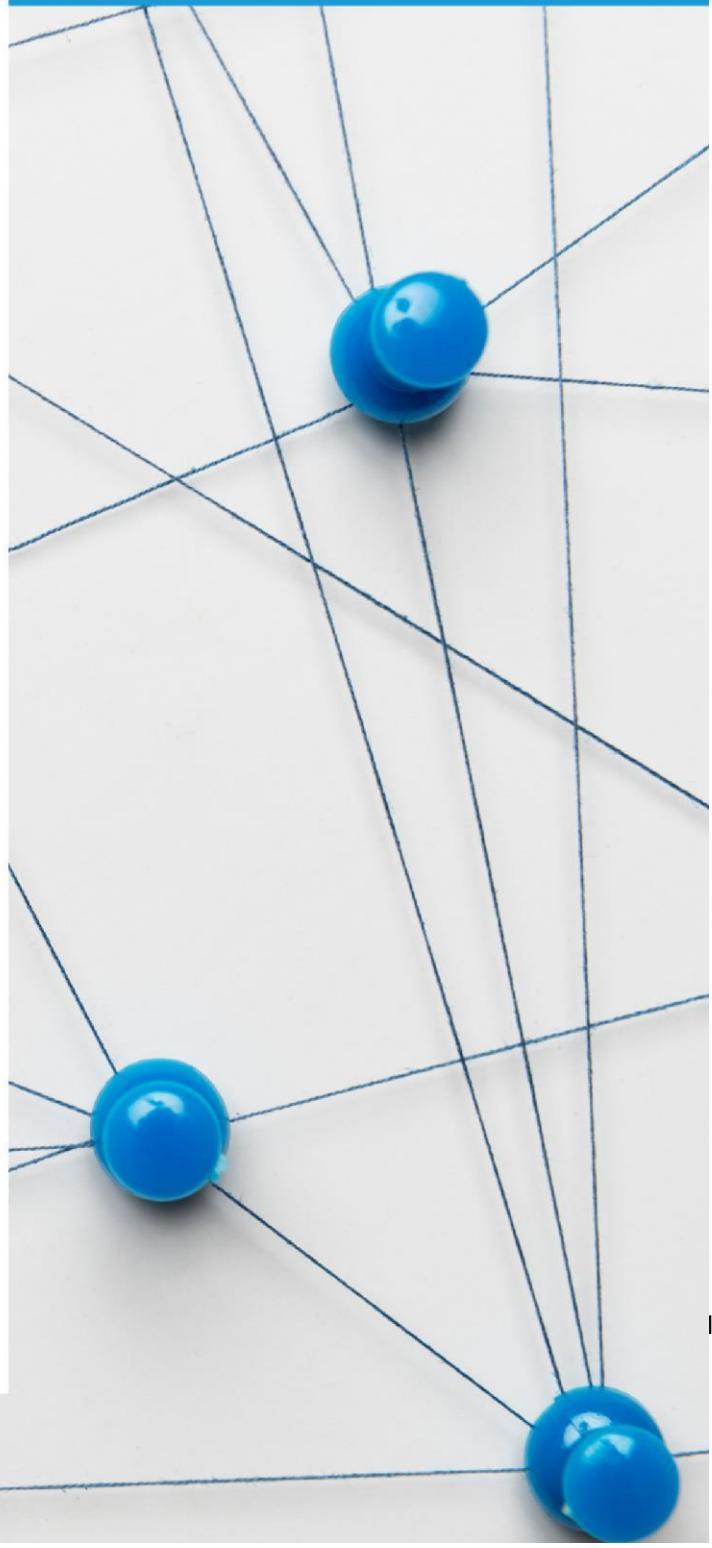


THE METHOD

CHAPTER 2: Systems Thinking and Behaviour

By Dr Stefan Kaufman, Dr Denise Goodwin,
Jenni Downes, Dr Peter Slattery.

www.behaviourworksaustralia.org



INTRODUCTION

How can a country encourage citizens to recycle correctly? In response to the 2017-18 'waste crisis'¹ in Australia, Amy Arbery's Behavioural Analysis Team in the federal government's environment department were asked to provide advice for the crossjurisdictional working group. The group was developing an action plan for the National Waste Policy 2018, which included a focus on how to reduce contamination and encourage correct household kerbside recycling. Recognising a need for improving education and awareness activities at all levels of government, and the potential of a national campaign, many stakeholders were arguing for more community education. However, from a systems and behaviour perspective, this raised some potential issues:

- *A focus on education implicitly assumes the primary problem is that Australian households lack the knowledge, skills and/or motivation to correctly sort the material, and keep problematic items out.*
- *Even if this is partly correct, education may be necessary, but not sufficient to solve the problem*
- *A campaign at a national scale risks assuming that the context of behaviour is the same throughout Australia.*

Indeed, our research found that Australia has collectively and unintentionally made it very hard for even highly motivated, knowledgeable and capable people to do the right thing when it comes to household recycling, wherever they are. To provide good advice, Amy's team needed to clearly define the problem, establish where household recycling behaviours sit within it, and identify what factors might be influencing waste outcomes. Conveniently, they had just helped form an initiative linking state and federal governments with behavioural researchers - the BehaviourWorks Australia [Waste and Circular Economy Collaboration](#). Together, we approached this challenge using systems thinking to understand the different systems recycling behaviours interact with, and used this to inform our advice and a shared agenda for developing and testing behavioural public policy interventions. This chapter uses the example of the waste crisis to show a number of ways systems thinking helps tackle behaviour change challenges.

SYSTEMS THINKING AND BEHAVIOUR CHANGE

Systems thinking is 'the art and science of making reliable inferences about (system) behaviour by developing an increasingly deep understanding of underlying structure' (Richmond, 1994). Systems analyses are often motivated by valued (or problematic) emergent

¹ Until late 2017, Australia, like many high-income countries, exported mixed household recyclables (plastics, glass, paper, cardboard etc.) predominantly to China. Citing local health and environment impacts, and poor economic value, China's import standard for acceptable contamination reduced to 0.5%. At the time in Australia, ~10-30% by weight of kerbside recycling was 'contamination' (i.e., soft plastics, clothes, rotting food, nappies, composite packaging etc.). See Jenni Downes' article outlining the issues in [The Conversation](#), and our [rapid evidence and practice review](#).

properties of systems— for example health and wellbeing, ecosystem function, prosperity, and, yes, the quality of recycling and its consequences.

While systems thinking ‘simply’ involves analysing a situation or problem of interest, and representing it as a series of elements or parts and their interacting relationships, it can be challenging. Systems literature, and practice, is laden with jargon and specialist techniques which can make it seem harder than it is. Even if it is not always easy, it’s very useful. In particular, ‘wicked’ policy problems are inherently persistent and recurring - their

complex root causes defy narrow problem definitions and simplistic responses. It is typical of wicked policy problems that multiple agencies and players may hold parts of the problem, solutions, and mandate to act on it. They are very difficult to clearly define or resolve without a systems thinking perspective, and examining the problem through the lens of behaviour in context can help make the situation, and opportunities to improve it, clearer for all involved.

For instance, an initial step in our work with the Waste and Circular Economy Collaboration involved working with Amy’s team to conduct a rapid evidence and practice review (see [Chapter 1](#) for more information on evidence reviews) and stakeholder workshop of government, community and business groups. Figure 1 summarises some of the drivers of misunderstanding and confusion thought to contribute to incorrect recycling behaviours that emerged from the review and workshop. Translated into systems thinking, these themes highlight that integrated solutions are required, many of which involve making correct recycling the path of least resistance to achieve desirable waste outcomes (Kaufman et al., 2020).

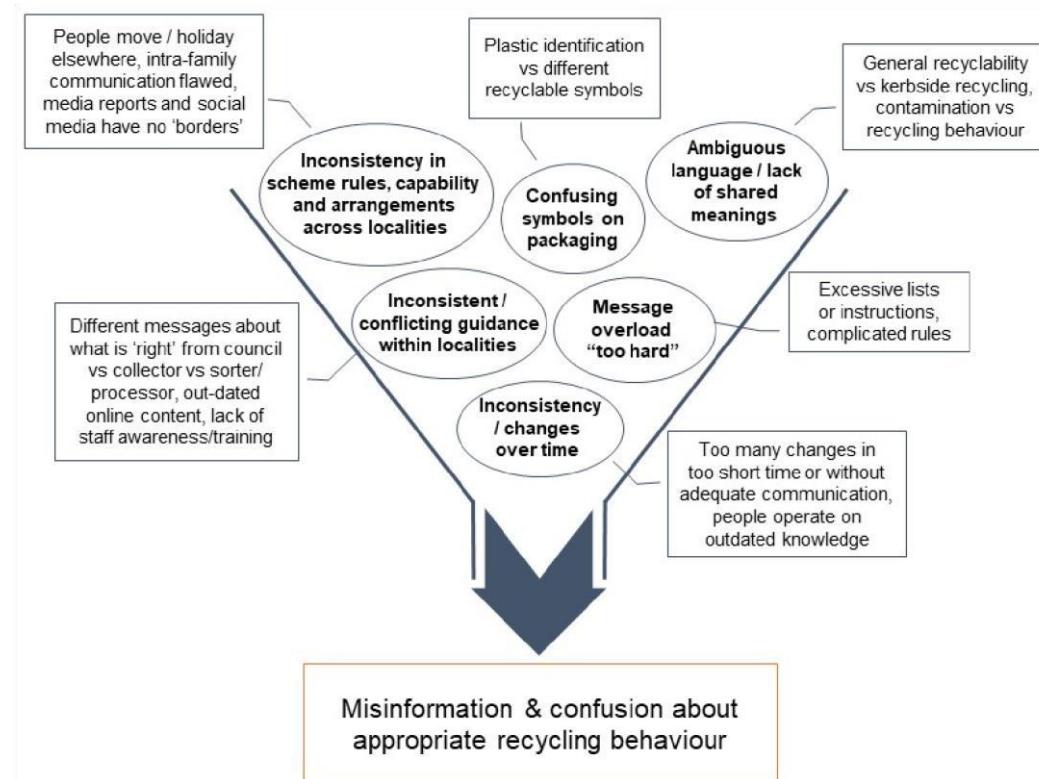


Figure 1: Some of the contributors to misinformation and confusion about correct recycling behaviour. As Figure 1 shows, the evidence suggested that a range of ‘upstream’ factors influence peoples understanding, and those understandings themselves reflect broader arrangements across economic, social, technical and biophysical systems. Systems thinking can help us to explicitly take into account how broader contextual factors influence human behaviour, integrate different information and views, and build a mutual understanding and shared response agenda to a problem (Brown et al., 2010; Vayda, 1996).

This is one of the ways systems thinking can highlight where integrated solutions are required that involve making changes at different levels in the system. Without the system thinking element, we likely would have focused on the real, but limited, opportunities to improve the quality of kerbside recycling through behavioural interventions. But recycling behaviours are constrained by, and contribute to, entities and relationships well beyond the boundaries of the household. So, instead of focusing only on the ‘presenting problem’ of the waste crisis and the contributions of household recycling behaviours, using systems thinking helped partners in the Waste and Circular Economy Collaboration identify and agree on the need for behavioural public policy experiments across the waste system (see Figure 2).

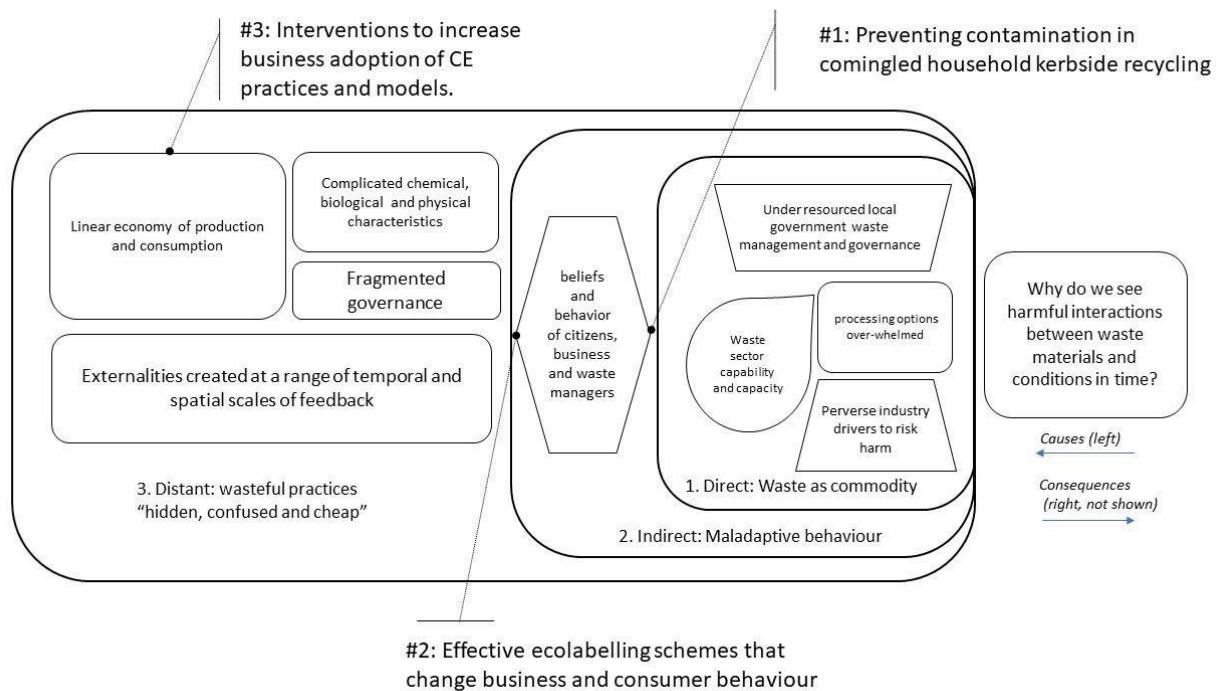


Figure 2: Using systems thinking via causal (and effect) mapping to frame a behavioural public policy research collaboration².

This example highlights that when we consider a given policy problem using systems thinking, we can helpfully translate our understanding of it into the following generalisable terms and concepts:

A system is composed of discernible parts/elements/agents that interact to constrain each other’s behaviour. Mutual constraints, operating between the parts of the system, limit the range of behaviours available to the system as a whole, and delimit

its boundaries – defining its ‘emergent’ (or synergistic) properties. System behaviour arises from the internally-generated forces imposed on parts of the system by other parts of the system – i.e., within and across different nested hierarchical scales of organisation over time, space and levels of complexity (Dyball et al., 2005; Hirschheim, 1983; Meadows and Wright, 2008; Richmond, 1994).

² This diagram draws on a much more detailed system map produced by Clarke (2018).

How do systems thinking and behaviour change approaches complement each other?

Applying a whole-of system approach to a problem generates a rich understanding of the issues and numerous associated behaviours, but can be less directly influential on the behaviours themselves. This is because systems thinking encourages us to consider organic, integrated wholes, composed of interrelated parts and subsystems. It highlights how change may be needed in multiple parts, and how improvements in one area may have unintended negative consequences in another. It's not always clear where or how rapid, substantial intentional change can be 'made' to happen, nor even how to change any one element without engaging with many (although analytical principles discussed in the 'do it yourself' example below help).

Conversely, focusing in on behaviours can be incredibly insightful and impactful for understanding and influencing a specific behaviour by a given population, but can cause problems if done too 'early' – i.e., without the contextual understanding of the causes and consequences of behaviour at a system level. Behavioural science tends to encourage us to apply powerful analytical 'reductionist' thinking to parts of the system, and attempts to identify the complicated but identifiable elements that need to be influenced to change individual behaviour. This is typically done through a carefully crafted intervention, based on a deep understanding of the reasons why people do or do not enact a specific behaviour in a time and place. However, this risks excluding broader perspectives and evidence, downplaying interactions with contextual factors, and the unintended consequences of intervening.

These contrasts suggest that both systems thinking and focused behaviour change³ techniques are needed at various stages of behavioural public policy to balance the advantages and risks of both. Human behaviour is caused by, and effects, multiple interacting systems (e.g. social, economic and environmental). The more complex or 'wicked' a policy problem is, the more critical it is that you consider systems when seeking behaviour change. In short, just focusing on a single behaviour or outcome within a small slice of the system can lead to unintended and undesired outcomes, problem mis-identification, and potentially reinforce or exacerbate the starting problem, or create others just as bad or worse⁴.

Recognising this, BehaviourWorks uses systems thinking most explicitly in the ‘Exploration’ phase of The Method: to consider the whole of an issue, identify the different parts and players involved with the issue and consider how they are related. This often helps reframe our understanding of the problem, and ‘whose behaviour could change how’, and therefore the intervention and behaviour change options. This is important pragmatically also, because a lot of attention and resources are focused on just one or a few behaviour changes for most of the later stages of The Method (Deep Dive, Application). However, systems thinking is still useful in these later stages, particularly when considering intervention implementation and scaling interventions from a successful trial (Best and Holmes, 2010).

Tools facilitating systems thinking are diverse, from quick workshop exercises, to extended stakeholder engagement, modelling and more. The specific tool we introduce here – cause and effect mapping - is just one of many systems tools useful in behavioural public policy and beyond.

³ That are also integrative and interdisciplinary, not locked into narrow models of human behaviour.

⁴ See this [pre-print of a recently submitted journal paper](#) for a longer discussion of these issues in the context of socio-technical transitions and behaviour change (Kaufman et al 2021).

USING CAUSE AND EFFECT MAPPING

So how do you start? People building sophisticated models frequently start with pen, paper and sticky-notes, and you can too. The process outlined below (with much iteration over three workshops and a desktop review) was used to produce the maps framing waste behaviour change interventions shown in Figure 2 .

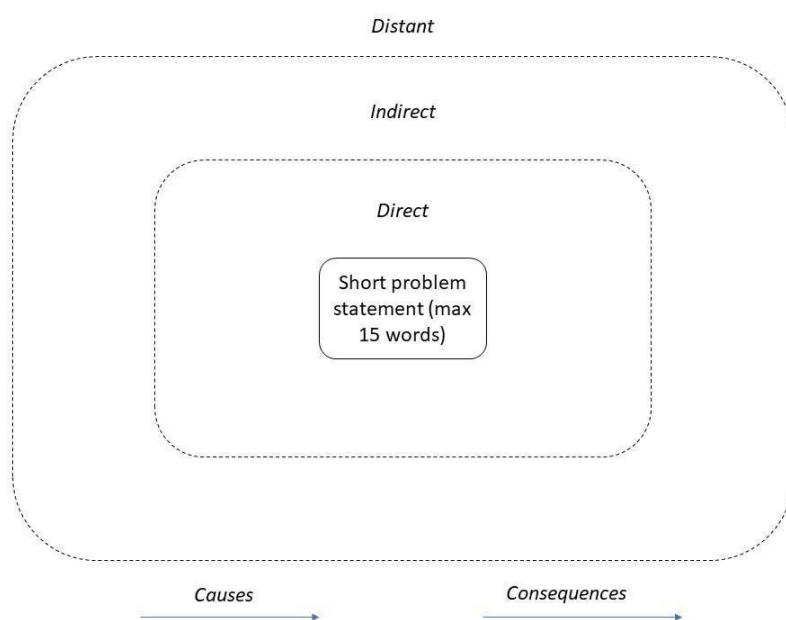


Figure 3: A simple cause and effect system map template

1. Sketch the above diagram (Figure 3) on a large piece of paper, showing a problem, and expanding bands of cause and effect on the problem.
2. Write the central situation, problem or opportunity in a short statement, simply and factually stating the subject of interest: i.e. ‘Two thirds of Australians are overweight or obese’.
3. Brainstorm different reasons you think this problem exists, and list them individually on small sticky notes, as you will want to move them around as you start the next step (one issue/reason per note).
4. Place them in the diagram based on how directly or indirectly you think they contribute to the problem. Is it a cause of the problem? Place it to the left. A consequence? To the right. If it appears to be both, you may want to break it down into two related, but different elements, and connect with arrows (see 6.).
5. Channel your inner child by asking repeatedly ‘why (does this problem exist)?’, with each question stepping back along chains of entities and relationships, adding elements as they occur. We use three ‘bands’ of causation above, but simple proximity works too.
6. In addition to proximity to the problem, how are these elements inter-related?
 - a. Illustrate relationships by adding arrows showing how you think one element influences another, with the thickness of the arrow showing how strong you think this influence is.
 - b. Each connection gets its own arrow (no single double-headed arrows! – this can misrepresent two way influences as equal).
 - c. Consider if the link is positive (+) or negative (-) – so a simple positive loop would get two arrows and a (+) sign.
7. Take a break. Looking again, see if you can identify any of the following relationships and features:
 - a. ‘Causal chains’ of elements and relationships that influence the problem. Sometimes targeting behaviour change at or near the root of the chain is much better value for effort.
 - b. ‘Cat’s claws’ – elements with many arrows leading into them. Think of a cat’s claw in a jumper – you need to unpick each one (driver) to remove it. Do not target them lightly!
 - c. ‘Octopus arms’ – sometimes an element influences many others, and could be high value for effort to target.
 - d. ‘Here be dragons’/‘call a friend’ – elements will appear on your map that seem well out of your ability to influence or mandate to act.⁵ Consider which ones need to be recognised as ‘too hard/risky’, or ‘parked for now’. But also consider whether partnering with others, who can more legitimately, efficiently and/or effectively target that element, is wiser.
 - e. ‘Fact, or opinion’ – consider which elements and relationships you know a lot about, which ones you might only think you do, and ones which need more evidence and analysis. This is a step where fresh or diverse eyes and minds can be helpful.

After 2 hours at this, you should be in a much better position to answer questions like:

- What is the problem, really?
- What is the context of the problem? What do we need to know more about?
- What are the important elements contributing to, and flowing from, the problem?
- Whose behaviour could change to improve the situation?
- What intended and unintended consequences from intervening can we anticipate?

ADVANCED APPLICATIONS

While the above is relatively easy on your own or in a small group, systems maps are representations, not reality, and there is the risk of embedding existing ‘group think’, assumptions, power dynamics, ‘false’ certainty, incorrect beliefs and biases. Experts can help reduce this risk by being ‘critical friends’ and applying evidence informed reflection and peer review through, for example:

- Facilitation and integrating diverse perspectives – people can reasonably disagree about complex situations, and indeed unpacking the conflict can be very insightful.
- Knowledge translation of rigorous, quality assured inputs like evidence reviews, data analytics, research.
- Rigorous and internally consistent systems models. Sophisticated methods and analytical tools abound. For example, agent based modelling of behaviour change interventions (Hansen et al., 2019; Schlüter et al., 2017).
- Methods and experience in managing, communicating and interpreting systems maps - they get unwieldy and dense quickly.
- Translating the map to ‘whose behaviour could change how’ and where behaviour change ‘sits’ (as a cause and consequence) within the problem (see later Method Book chapters).

⁵ Remember it is typical of wicked policy problems that multiple agencies and players may hold parts of the problem, solutions, and mandate to act on it. Looking at behaviour in context can help make this clearer, and therefore the case for working together.

Indeed, mapping exercises can be enriched by a range of useful frameworks that prompt thinking about broader systems. This potentially engages disciplines across the natural and social sciences and humanities (Jackson, 2018). Linking different systems and behaviour can be both insightful, and practical. Consider for example how

BehaviourWorks Australia’s Jenni Downes’ research maps the systems co-evolving with recycling behaviours (see Figure 4). Integrating published research and applied behavioural policy tools in systems workshops with waste educators, she detailed a wide range of possible influences to consider and use in planning interventions. This can inform subsequent steps in The Method.

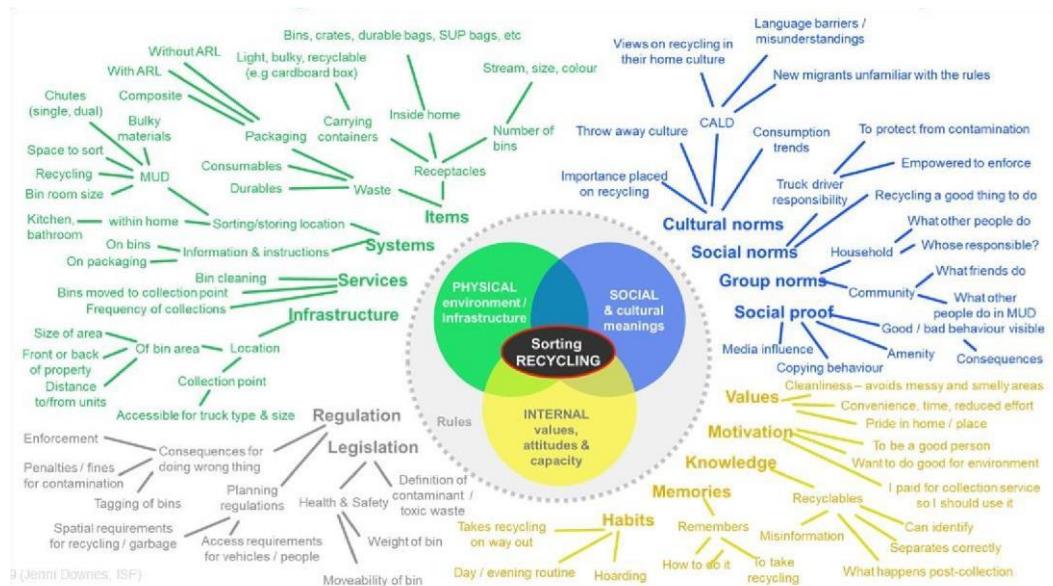


Figure 4: The systems causing, and effected by, recycling behaviour

Indeed, a range of frameworks situate behaviour in diverse policy contexts and explore the implications for changing individuals' behaviour, and the systems they are participating in – racism, health, sustainable consumption, energy, socio-ecological change, socio-technical transitions and more⁶. Systems thinking applied to behaviour change is a valuable way of bringing together diverse sources and types of evidence and knowledge, and grounding it in the practicalities of how do we help people perform a behaviour differently.

CONCLUSION

In the end, Amy's team were able to support the development of the action plan for the National Waste Policy 2018² with advice incorporating systems perspectives. They underlined the overall importance of changes beyond the individual to make correct recycling easier. In particular, coordinating local, regional and national initiatives, and transforming systems that shape consumer behaviour. Taking a systemic approach to behavioural public policy, Amy and colleagues drew on inputs across The Method's Exploration phase, including a rapid evidence and practice review ([see Chapter 1](#)) on recycling contamination, a national stakeholder workshop, the

⁶ See for example: Akenji, Lewis, Huizhen, and Chen, 2016; Bruckmeier and Pires, 2018; Darnton 2013; Feagin, 2013; Kaufman et al. 2021a, Kaufman et al. 2021b, Public Health England, 2018; Schell et al., 2020; Vayda, 1996.

waste causal influence system map, and the shared knowledge and advice of staff and partners within the Waste and Circular Economy Collaboration. Overall, this work supported the need for policy initiatives reforming systems that frame recycling and sustainable consumption behaviour in Australia, including product design and labelling, business innovation to offer 'circular economy' products and services, local campaigns, and improved

² [National Waste Policy Action plan](#), p. 10.

collecting and sorting systems, bin designs and more to support correct kerbside recycling behaviours (see [BehaviourWorks Australia collaboration publications and products](#)).

REFERENCES

- Akenji, Lewis, Huizhen, Chen, 2016. A framework for sustainable lifestyles: Determinants and strategies (No. DTI/1717/PA). United Nations Environment Programme.
- Best, A., Holmes, B., 2010. Systems thinking, knowledge and action: towards better models and methods. *Evid. Policy J. Res. Debate Pract.* 6, 145–159. <https://doi.org/10.1332/174426410X502284>
- Brown, V.A., Harris, J.A., Russell, J.Y. (Eds.), 2010. Tackling wicked problems through the transdisciplinary imagination. Earthscan, London; Washington, DC.
- Bruckmeier, K., Pires, I., 2018. Innovation as Transformation: Integrating the Socio-ecological Perspectives of Resilience and Sustainability, in: Pinto, H., Noronha, T., Vaz, E. (Eds.), Resilience and Regional Dynamics, Advances in Spatial Science. Springer International Publishing, Cham, pp. 209–231. https://doi.org/10.1007/9783-319-95135-5_11
- Clarke, J., 2018. EPA problem solving workshop report: Waste. Unpublished report produced by Minds at Work on behalf of the Victorian Government Department of Land, Water, Planning and Environment, EPA Victoria and Sustainability Victoria, Melbourne, Australia.
- Darnton, A., Evans, D., Scotland, Scottish Government, 2013. Influencing behaviours: a technical guide to the ISM tool. Scottish Government, Edinburgh.
- Dyball, R., Beavis, S., Kaufman, S., 2005. Complex adaptive systems: constructing mental models, in: Keen, M., Brown, V.A., Dyball, R. (Eds.), Social Learning in Environmental Management: Towards a Sustainable Future. Earthscan, London; Sterling, VA, pp. 41–59.
- Feagin, J., 2013. Systemic Racism: A Theory of Oppression. Routledge, New York.
- Hansen, P., Liu, X., Morrison, G.M., 2019. Agent-based modelling and socio-technical energy transitions: A systematic literature review. *Energy Res. Soc. Sci.* 49, 41–52. <https://doi.org/10.1016/j.erss.2018.10.021>
- Hirschheim, R.A., 1983. Assessing participative systems design: Some conclusions from an exploratory study. *Inf. Manage.* 6, 317–327. [https://doi.org/10.1016/0378-7206\(83\)90040-X](https://doi.org/10.1016/0378-7206(83)90040-X)
- Jackson, M., 2018. Systems Research and Behavioral Science Journal [WWW Document]. URL <https://onlinelibrary.wiley.com/journal/10991743a> (accessed 10.31.18).
- Kaufman, Stefan; Saeri, Alexander; Smith, Liam; Malek Pour, Shirin; Raven, Rob (2021a): Kaufman et al 2021
Behaviour change and sustainability transitions.pdf. Monash University. Unpublished pre-print submitted to Environmental Innovation and Socio-technical transitions (29/01/2021). <https://doi.org/10.26180/13654502>

Kaufman, Stefan; Kirchherr, Julian; Curtis, Jim; Downes, Jenni; Saeri, Alexander; Slattery, Peter; et al. (2021b): Barriers to the Circular Economy in Australia and the European Union (EU): A comparative mixed methods review. Monash University. Unpublished pre-print under review with the Journal of Cleaner Production. <https://doi.org/10.26180/13654511>

Kaufman, S., Meis-Harris, J., Spano, M., Downes, J., 2020. What ‘works’ behaviorally to reduce contamination of recycling at the kerbside (a rapid evidence and practice review). BWA Waste and CE collaboration, BehaviourWorks Australia, Monash Sustainable Development Institute, Monash University, Melbourne, Australia.

Meadows, D.H., Wright, D., 2008. Thinking in systems: a primer. Chelsea Green Pub, White River Junction, Vt.

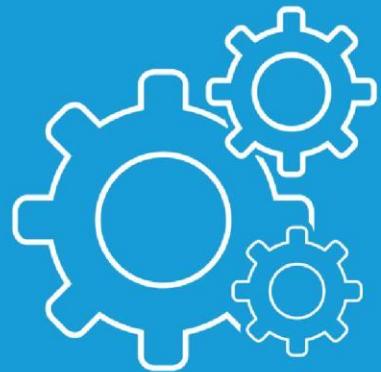
Public Health England, 2018. Improving people’s health: Applying behavioural and social sciences. Department of Health, London.

Richmond, B., 1994. Systems thinking/system dynamics: Let’s just get on with it. *Syst. Dyn. Rev.* 10, 135–157. <https://doi.org/10.1002/sdr.4260100204>

Schell, C.J., Dyson, K., Fuentes, T.L., Roches, S.D., Harris, N.C., Miller, D.S., Woelfle-Erskine, C.A., Lambert, M.R., 2020. The ecological and evolutionary consequences of systemic racism in urban environments. *Science* 369. <https://doi.org/10.1126/science.aay4497>

Schlüter, M., Baeza, A., Dressler, G., Frank, K., Groeneveld, J., Jager, W., Janssen, M.A., McAllister, R.R.J., Müller, B., Orach, K., Schwarz, N., Wijermans, N., 2017. A framework for mapping and comparing behavioural theories in models of social-ecological systems. *Ecol. Econ.* 131, 21–35. <https://doi.org/10.1016/j.ecolecon.2016.08.008>

Vayda, A.P., 1996. Methods and Explanations in the Study of Human Actions and their Environmental Effects. Centre for International Forestry Research and World Wide Fund for Nature, Jakarta, Indonesia.

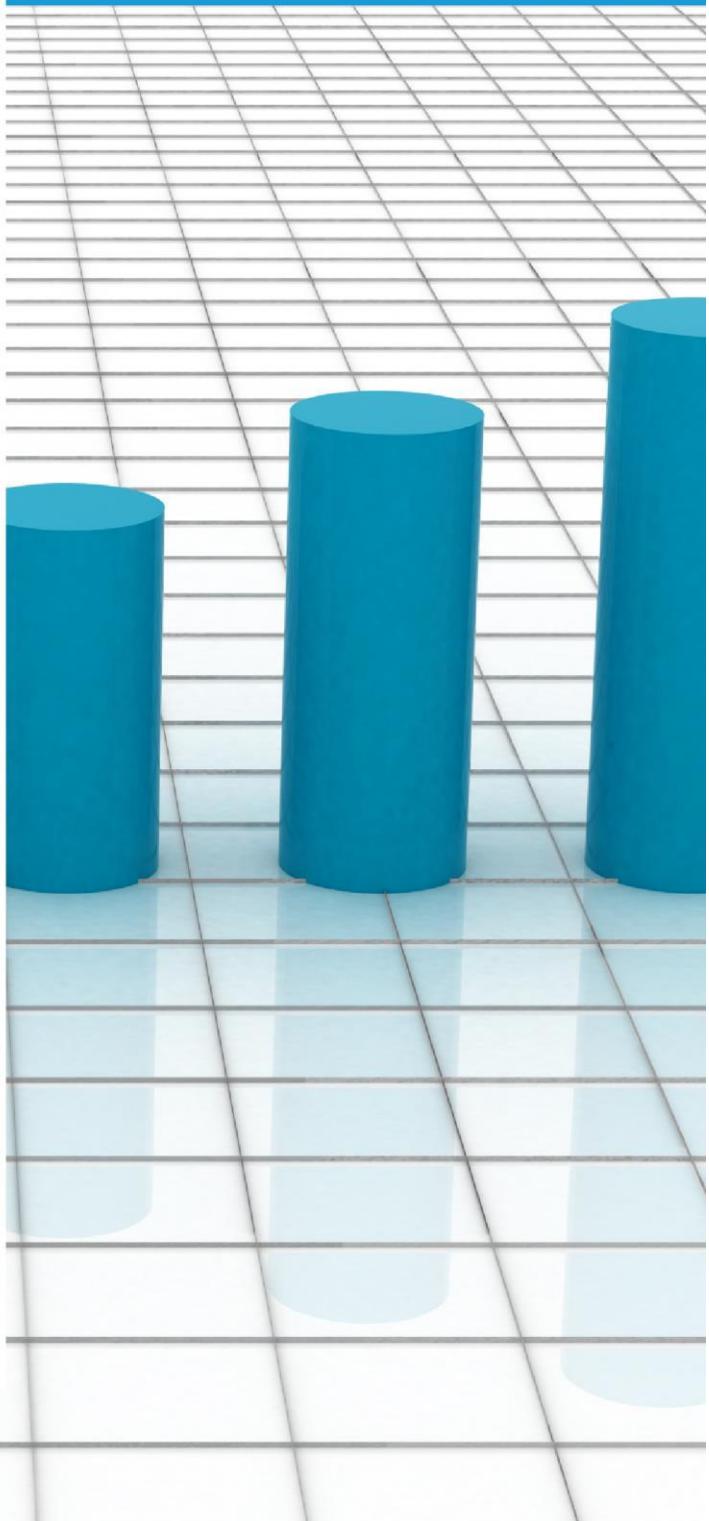


THE METHOD

CHAPTER 3: Exploring the problem with data

By Kim Borg, Dr Kun Zhao and Dr Fraser Tull.

www.behaviourworksaustralia.org



INTRODUCTION

Sarah is a regional manager at a large charity organisation. She oversees a large number of retail stores selling quality second-hand clothing and other household goods donated from the community. Unfortunately, not all the donated goods are suitable for sale and must be sent to landfill at the cost to the organisation. To make matters worse, some people don't donate their items appropriately during store hours, instead leaving them outside the store overnight where the best items are often stolen or become weathered and unfit for sale. Other people take advantage of the charity and dump waste outside the store to avoid having to dispose of it appropriately. This costs the organisation money which directly impacts how much funding is available for community programs. Sarah wants to stop people from dumping their unwanted goods and waste. She decides to develop a behaviour change intervention to tackle this costly issue. Before making any assumptions, Sarah decides to look at the data she has available to learn more about the problem.

While evidence reviews are a great way of synthesising published knowledge (Chapter 1) and systems thinking helps to generate a rich understanding of an issue within a wider complex network (Chapter 2), one of the best ways of understanding a specific problem in your particular context and in detail is to look at your data. Numbers provide a rigorous means of identifying *who* needs to do *what* *when* and *where*.

In this chapter we introduce some common methods we use when analysing data in order to better understand a collected as a source of information" (Australian Bureau of Statistics, 2020). Data can be *quantitative* (i.e. numeric information), which may be collected routinely or purposefully measured through surveys, polls, and observations. Data may also be *qualitative* (i.e. nonnumeric, descriptive, and often language-based). This chapter focuses on methods for analysing quantitative data in order to explore a problem. However, quantitative and qualitative data analysis can also be applied in other phases throughout The BWA Method (see Box 1).

Throughout this chapter, we cover *simple analysis techniques* (e.g. descriptive statistics, such as averages and frequency counts) and *complex analysis techniques* (e.g. inferential statistics, such as modelling). We also discuss approaches for examining differences within your audience through *segmentation* (e.g. cluster analysis). This chapter focuses on describing why and when you would use these techniques in behaviour change practice through real-world applications. While it is not intended to be a statistics how-to guide, there

problem. Broadly speaking, data refer to "measurements or

observations that are Box 1. Where does data analysis fit in The BWA Method?

Data analysis is particularly versatile because it can be used for different purposes across all three phases of The BWA Method (*Exploration*, *Deep Dive*, and *Application*). For this reason, there is no single "segment" dedicated to data analysis. At different phases, data analysis can be used in multiple ways, including:

Understanding the problem by determining which behaviours and people are contributing to the outcome of interest (*Exploration*: *Review and Collect Evidence on the Problem*; *Identify Priority Behaviours*);

Identifying factors that are influencing behaviours and audiences (*Deep Dive*: *Review and Collect Evidence on What Works*; *Understand Drivers and Barriers*) and;

Determining if your intervention made an impact and which groups were more or less impacted (*Application*: *Trial the Interventions*).

This chapter will focus on examples of data analysis in the Exploration phase of The BWA Method.

are some basic concepts and terms that will help to understand why different types of data analyses are useful – see the glossary in Table 1.

Table 1: Glossary of basic statistical concepts in this chapter

TERM	Description
Variable	A characteristic that can be measured or counted. Examples of a variable include age, income, whether someone received a vaccine, amount of illegal waste, number of hours worked, and selfreported attitude.
Predictor (or independent) variable	A variable that is thought to explain changes in an outcome variable – e.g. temperature (predictor variable) helps explain the number of ice-creams sold.
Outcome (or dependent) variable	A variable that is thought to change as a result of variation in the predictor variable – e.g. the number of ice-creams sold (outcome variable) depends on how warm or cold the temperature is.
Continuous variable	A variable which can take an infinite number of values between two points of measurement. For example, height is a continuous variable as any two measurements could have an unlimited number of values between them.
Categorical variable	A variable which can take a finite number of values or categories. For example, school grade is categorical as it can take only one of 6 values (A, B, C, D, E, F).
Dichotomous variable	A variable which can take on one of two values. For example, whether someone has returned a consent form (yes or no).
Descriptive statistics	Statistics that summarise the basic properties of a given dataset. These can include the mean, frequency counts, standard deviation, and range.
Inferential statistics	Statistics that allow broader conclusions or inferences to be drawn from the sample, particularly whether the findings in the dataset represent true relationships or if they are likely to have occurred by chance. These include ttests, correlations, and regressions. Inferential statistics are also the basis for predictions and estimates.
Correlation	The extent to which two variables are associated. Two variables can be <i>positively</i> correlated (as one goes up, the other also tends to go up), <i>negatively</i> correlated (as one goes up, the other tends to go down), or show <i>no correlation</i> (as one goes up, it does not tell us anything about what the other will do).
Segmentation	The process of dividing different units (e.g. people, shops, schools, hospitals) into groups that are more alike to each other than to members of other groups on a given characteristic(s).

GETTING THE MOST OUT OF YOUR DATA

Often, we assume we know the scale or cause of a problem. However, our subjective understanding of the issue may be inaccurate or biased. Numbers have a great way of taking assumptions out of the equation.

One of the first steps we do in any behaviour change project is to ask “do you have any data on the issue?” or “do you need to collect some data to understand the issue?” From our experience, government and industry partners often have more data than they think, or do

not always utilise their data in a way that really helps them understand a problem or identify who and what needs to change.

Data analysis, put simply, is examining and testing data in order to gather useful insights. Data can provide information about the scale of a problem (e.g. the amount of waste left at each store per week), as well as characteristics about different types of cases (e.g. size of the store, operating hours, distance to car park, etc.). At an individual level, it can also include socio-demographic information about your audience (e.g. age, income, education level, etc.), as well as more complex information (e.g. personality traits, preferences, attitudes, or values). When applied appropriately, data analysis can give us a clear and indepth picture of what is going on, as well as direct us to how we might fix the issue.

Simple data analyses

Simple analytical techniques include **descriptive statistics**, which summarise the properties of your dataset. These include statistics like the **mean** (i.e. what's the average score among a group of people) and **frequency counts** (i.e. how many people fall into different categories). The purpose of descriptive statistics is to simplify a large amount of information into a summary format that's easier to interpret.

Descriptive statistics are useful for determining what the problem is and where it is occurring – or if there is a problem at all. Descriptive statistics are most useful when you want to get an overall picture of what's going on or see what's happening on “average”. For example, instead of being bombarded with various amounts of waste at each of the 50 individual stores, Sarah could calculate the average (mean) by adding the total amount of waste collected overall and dividing by 50, to get an overall picture of how much waste is being accumulated per store. Furthermore, Sarah could use descriptive statistics to examine the frequency distribution of the amount of waste across all 50 stores. This might tell her whether the problem is experienced to the same extent for all stores (i.e. most stores experience similar levels of waste) or whether the problem is driven by a severe amount of waste in a few select stores while most others are not impacted.

Complex data analyses

Complex analytical techniques include **inferential statistics**, such as **modelling**. Modelling refers to a set of analyses which explore the relationship between one or more predictor (independent) variables (e.g. the size of the charity store, or the economic profile of the area) and an outcome (dependent) variable (e.g. the amount of waste dumped at the store). When using these approaches, we're trying to identify relationships between predictors and outcomes. In doing this we can see what factors are potentially increasing or minimising the problem, and therefore what needs to change.

Where descriptive statistics summarise the properties of the dataset that you have in front of you, more complex analyses allow you to make inferences by concluding whether the findings in your dataset are likely to reflect real explanations, and not arbitrary results that were simply due to chance. To get reliable findings, most analyses depend on a set of assumptions (e.g. the way the data are distributed must fall in a particular pattern) and other requirements of the data (e.g. a sufficiently large number of people or observations in your dataset). Quantitative behavioural experts are trained in how to test for these assumptions,

what can be done when these assumptions are not met, and how to evaluate existing research for robustness based on these assumptions.

At a basic level you can look at relationships between two variables, or bivariate relationships – e.g. using **correlations** to explore the relationship between one predictor variable and an outcome variable (i.e. do larger stores have more waste compared with smaller stores? Or, do stores with a greater distance to their carpark have less waste compared to stores with a shorter distance?). You could also look at the differences in an outcome variable between two groups using a **t-test** for differences in means (e.g. do charity stores in low socioeconomic areas have more waste dumped on average than stores in high socio-economic areas?), or **chisquare** for differences in frequency distributions.

The real strength of modelling, however, comes when we explore the relationship between many predictor variables and an outcome variable, all at the same time. This can be done using **regression analyses**. When we enter multiple predictor variables into a model, it allows us to pinpoint exactly which predictors are having a unique effect on the outcome, while also controlling for many different relationships. For example, in a project which explored how Australians used or did not use the internet for various activities, we found that fewer older people and fewer people who were retired used the internet compared to younger people and people who were working. However, when controlling for the effect of multiple factors at once using a regression analysis, we

found that ‘work status’ (i.e. being retired) did not have a unique effect on internet use, but age (i.e. being older) did. In other words, being older was associated with reduced internet use, and it just so happened that many older people are also retired. In other words, regression models help to tease out the real relationships from more arbitrary ones.

There are different types of regression analyses that should be used when testing multiple predictor variables, depending on the properties of your data. For example, when the outcome you are interested in is *continuous* (e.g. amount of waste at charity store) a **multiple linear regression** analysis is best. When your outcome variable is *dichotomous* (e.g. people who use vs. people who do not use the internet) **logistic regression** analysis is required. And if your outcome variable is categorical (e.g. different typologies of people who use the internet - those who rarely use the internet at all, those who use it for socialising and entertainment, those who use it for work and information, and those who use it for everything

- see example below) **multinomial logistic regression** is required. If you’re not sure which analysis is right for your data, it might be worth consulting with a quantitative expert to make sure you’re using the right technique.

Segmentation

Another analytical approach is audience segmentation, which involves understanding differences within your audience of interest and then dividing your audience into groups whose members are more alike than they are to members of other segments (Grunig, 1989). Segmentation allows you to (a) better understand the motivations and drivers of your target behaviour, and based on this, (b) separate your audience into meaningful groups for more precise and targeted behaviour change interventions.

One of the biggest mistakes you can make as a behaviour change practitioner is to assume that your audience is like you, but it is also a mistake to assume that members of your audience are like each other. In addition to socio-demographic and contextual characteristics, your audience will most likely have different attitudes and personality traits which will affect how different people respond to behaviour change interventions (Alkış & Temizel, 2015).

Simple and complex segmentation techniques can be used to break up your data into different groups. Simpler segmentation methods include dividing your audience based on predefined demographic or contextual criteria, such as age, gender, or location (e.g. segmenting charity stores based on store size). You could then collect additional data—through surveys, interviews, observations—and examine the differences between these groups in relation to your target behaviour using both descriptive and inferential statistics. Alternatively, you could work backwards and create segments based on whether your audience undertakes or does not undertake the desired behaviour (e.g. compliers and noncompliers). Then you could examine the differences between these two groups across a variety of characteristics - such as attitudes, personality traits, and demographics. In both these examples, the segmentation is defined *a priori*, that is, you decide what characteristics to segment your data on.

At a complex level, we can use statistical techniques (e.g. **cluster analysis**) to segment our population based on behavioural and psychographic characteristics (e.g. segmenting people based on their online behaviours). Cluster analysis takes into account multiple factors, including categorical (e.g. male, female) and continuous (e.g. age in years) data, and uses a statistical procedure to maximise the similarity *within* each cluster while also maximising the difference *between* each cluster. Cluster analysis is particularly useful when you want to use more nuanced characteristics (such as individual attitudes or behaviours) to create groups. Unlike simpler segmentation methods where you define your segments beforehand, cluster analysis is data-driven and produces the segments based on the quantitative qualities of your data.

MATCHING RESEARCH QUESTIONS TO TECHNIQUES

Before undertaking any data analysis, we need to determine our research questions in order to decide which analytical technique to use (see Figure 1). In other words, what questions are we trying to answer with data?

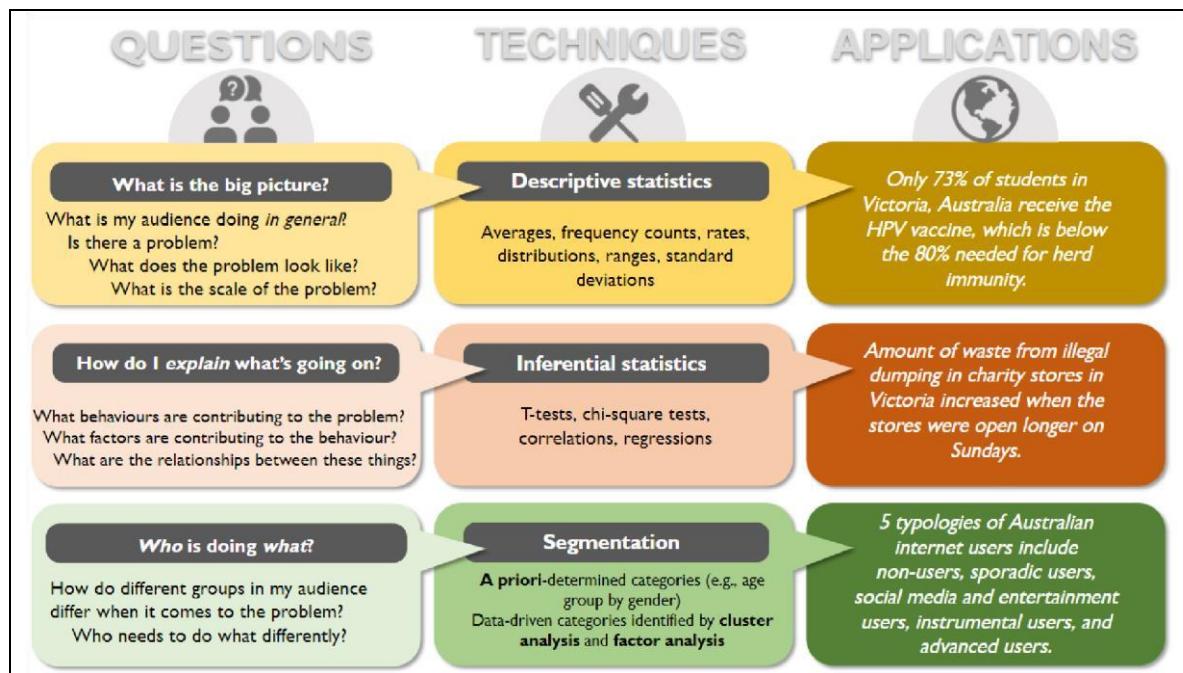


Figure 1: Matching research questions to data analysis techniques

What is the big picture?

In the first instance, your data can be used to paint an overall picture of what is going on and what your audience is doing. It may help to answer questions such as: *Is there a problem?* *What does the problem look like?* and *What is the scale of the problem?*

For example, using descriptive statistics, we looked at the proportion of Year 7 students who receive the full course of the Human Papillomavirus (HPV) vaccine which is administered as part of the Secondary School Vaccine Program in Victoria. It was found that every year, around 73% of Year 7 students received the full course of the vaccine. This is less than the proportion of the population required to be vaccinated to achieve herd immunity (target=80%) which leaves students vulnerable to preventable diseases. By simply looking at the proportion of students who did and did not receive the vaccination, we were able to identify a) that there *was a problem* (not enough students getting vaccinated to achieve herd immunity) and b) the *scale of the problem* (vaccination rates needed to increase from 73% to 80% or higher).

Similarly, in the case of charity store waste, Sarah can use descriptive statistics to find out the average (i.e. mean) amount of waste which comes from different sources. Across the stores, the mean amount of poor-quality donations is 11.41m³, but the mean amount of waste from illegally dumping items is even greater, at 16.63m³ (Wright, Smith, & Tull, 2018). Given this information, Sarah decided to focus her efforts on stopping people from illegally dumping their waste outside store operating hours.

How do I explain what's going on?

Your data could also help explain the different things that are contributing to your problem and identify the relationships between them. In the Exploration phase of The BWA Method, data can be used to answer questions such as: *What behaviours are contributing to the problem?* For example, in the case of low vaccination rates among Year 7 students,

descriptive statistics also provided information on the different behaviours contributing to the issue. A simple frequency count revealed that only 5% of Year 7 students received the HPV vaccine elsewhere. Therefore, low vaccination rates in the program were not just a matter of students using alternative services. Around 21% did not receive the vaccine because their parents declined consent and 31% didn't have a completed consent card. However, the real surprise was the percentage of students who didn't receive the vaccine simply because they were absent on the day of vaccinations (43%) – see Figure 2. In other words, nearly half of the students who didn't get vaccinated had a signed consent card, but weren't present to be vaccinated. Therefore, it was determined that attempts to increase vaccination rates should focus on getting students to attend school on immunisation day.

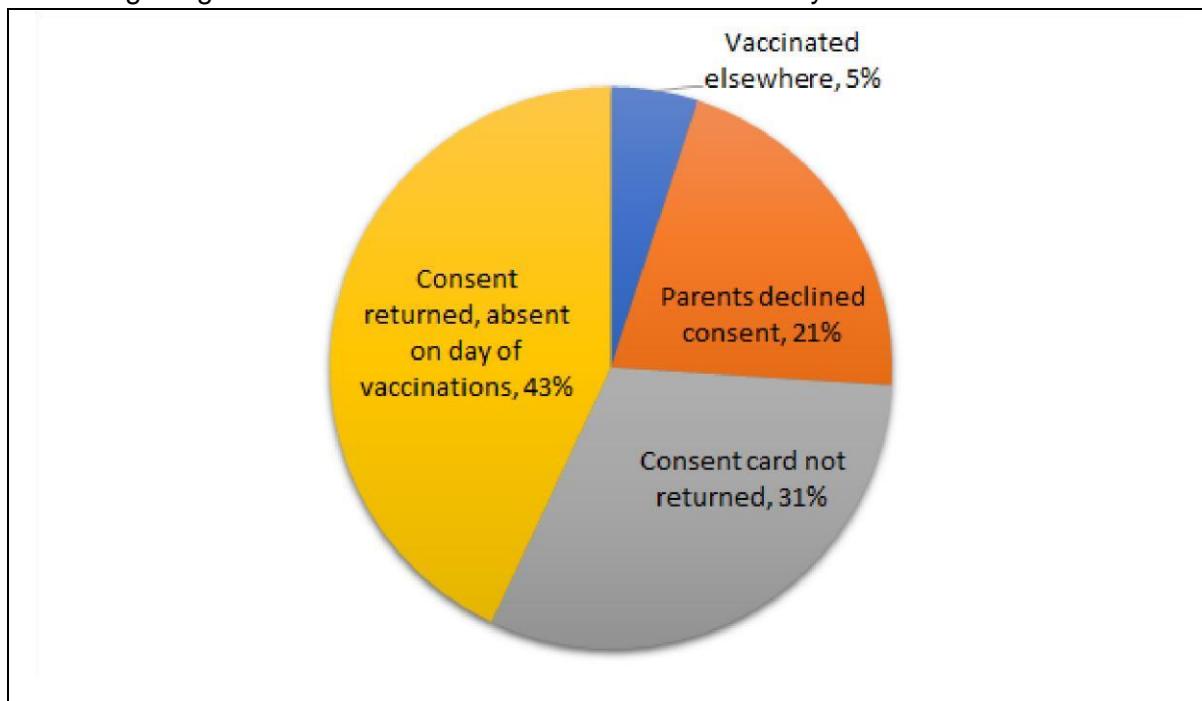


Figure 2: Behaviours associated with students not receiving the HPV vaccine within the school program

Data can also be used to delve further and answer questions such as: *What factors are contributing to the behaviour?* This can shed light on some of the drivers and barriers of behaviour, which is most useful during the Deep Dive phase of The Method. In the case of waste at charity stores, descriptive statistics already helped identify that illegally dumping waste outside of business hours was the biggest problem behaviour. Next, inferential statistics (modelling) was used to identify different factors that were associated with the amount of waste being dumped (i.e. the outcome variable).

Sarah decided to conduct a Multiple Linear Regression analysis to identify the factors associated with the amount of waste dumped at charity stores. After checking all of the assumptions, she ran the model and found a number of community level and site-specific factors that were associated with dumping, and a number of factors that were not associated with dumping (see Wright, Smith & Tull, 2018). Factors associated with increased waste were:

- Lower levels of education/occupation and greater mobility of residents in the area,
- Longer opening hours on Sundays,
- Having fencing around the store,

- Having bins/drop boxes brought in at night.

Meanwhile, distance to a landfill and distance to a carpark were not associated with increased waste when controlling for all the factors above. Armed with this information, Sarah knew that it was a combination of sitespecific characteristics and community-level characteristics that contributed to illegal dumping at charity stores. These findings also pointed to possible intervention options, such as improved maintenance and better placement of bins.

Who is doing what?

In the previous sections, we looked at how descriptive and inferential statistics can help shed light on the problem and the factors contributing to the problem in your audience as a whole. However, sometimes it's not enough to know what's going on overall. You may need to know: *how different groups in your audience differ when it comes to the problem?* Segmentation is a great way of identifying the 'who' in the overall question of *who* needs to do *what* differently, *when* and *where*.

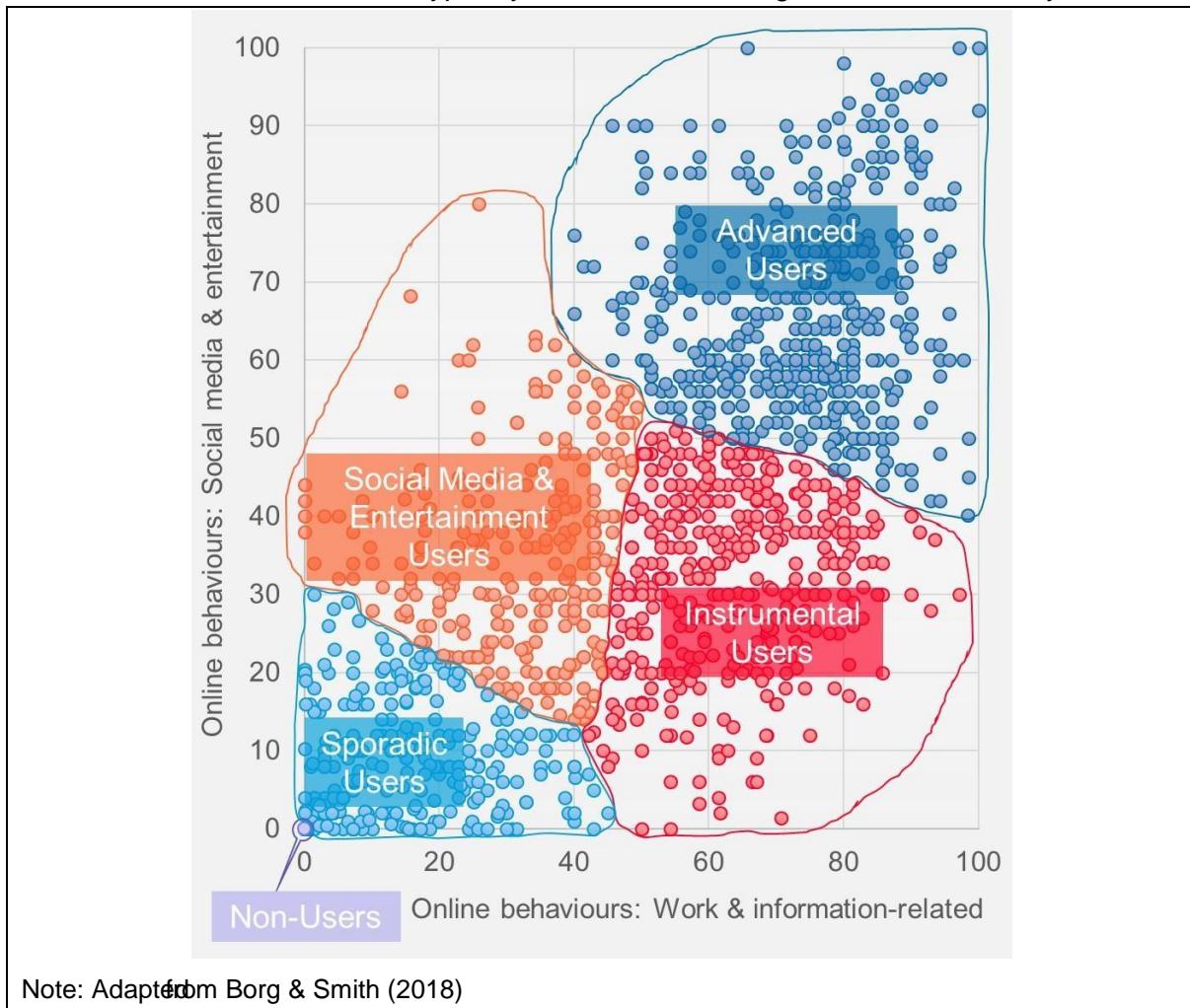
Segmentation is most powerful when you have the capacity to develop a targeted intervention, such as creating different messages that can reach different target audiences exclusively. However, it may not always be practical or necessary. In determining whether segmentation is appropriate, consider what information you have already and what you still need to know. It may also depend on what phase of The BWA Method you are in. During Exploration, we often want to know where the problem is occurring and who is contributing to it, so we might use measures of existing behaviour, contextual, and demographic criteria to identify the compliers and noncompliers. In the Deep Dive phase, we want to know what factors are influencing desirable and undesirable behaviours among different groups, so we might segment based on attitudes or other behavioural drivers. Finally, in the Application phase, segmentation can help to determine which groups were more or less impacted by a trial.

There are multiple criteria and strategies for segmenting your audience. Ultimately, it comes down to your research question, your problem, and existing knowledge about your target audience (e.g. what does the literature say about which audience segments are more likely to take up the target behaviour?). For example, in the study on internet behaviours described earlier, the research question was: *which Australians are most likely to be 'digitally excluded'?* We know from the literature that age and gender are associated with internet use in general. Based on this, we could have simply grouped people based on their age (18-44, 45-64, 65+) and gender (males and females) which would have produced six independent groups. We could have then used a combination of descriptive and inferential statistics to learn more about the online behaviours and attitudes of these groups.

However, we also know from the literature that digital exclusion isn't just about using and not using the internet. It's also about how people use the internet for different activities (e.g. communicating with friends and family, searching for jobs, playing games, etc.). We decided to conduct a telephone survey to collect data on how often Australians used the internet for a variety of activities. The types of online activities were grouped into two broad behavioural categories: (a) work and information-related behaviours and (b) social media and

entertainment behaviours. We then used cluster analysis to identify a *behavioural typology* using the two online behavioural categories as well as whether people identified as an internet ‘user’ or ‘non-user’. This resulted in five groups of internet users (see Figure 3) who engaged with the internet in different ways (Borg & Smith, 2018):

- *Non-Users*: never use the internet,
- *Sporadic Users*: rarely use the internet for both types of online behaviours,
- *Social Media & Entertainment Users*: prefer social-related online behaviours,
- *Instrumental Users*: largely prefer work and information-related behaviours,
- *Advanced Users*: typically use the internet, regardless of the activity.



Note: Adapted from Borg & Smith (2018)

Figure 3: Online behaviours of five clusters of internet users

Box 2. A note of caution about segmentation

By definition, segmentation involves breaking up a diverse and varied audience into segments which are definable, mutually exclusive, easily reachable, and large enough to have substantial effects (Grunig, 1989). Although this is crucial for the practical application of behaviour change interventions, it is important to be aware that this is sometimes at odds with underlying theory (and at times, evidence).

For example, categories used to segment can be arbitrary and exaggerate differences between groups. Consider two segments—a younger age group consisting of 18–34 year olds and an older age group consisting of 35–55 year olds. The difference between someone aged 34 and someone aged 35 is likely to be minimal compared to the difference between someone aged 18 and someone aged 55.

Likewise, most individual differences of a psychological nature (e.g. personality traits) are considered to vary by kind and along a continuum rather than falling into distinct categories or typologies in a binary fashion (Haslam, 2019). It is important to keep these limitations in mind when drawing conclusions or making recommendations based on the findings from your segmented data.

Once the groups were identified we then used descriptive and inferential statistics to examine and compare the groups in relation to their demographic (e.g. age, gender, location) and non-demographic (e.g. internet access, skills, and attitudes) characteristics. For example, people were more likely to be *Social Media & Entertainment Users* if they were female, did not have a tertiary qualification, and were concerned about online privacy. This process helped us get a better sense of which Australians were more likely to be ‘left behind’ in the digital age.

By knowing the online behaviours of each group, we also had useful information for developing future interventions. We knew that *Social Media & Entertainment Users* often use social media and online gaming platforms. Therefore, if we wanted to develop an intervention that promoted more ‘capital-enhancing’ online activities, like searching for job vacancies via job seeking websites, we could do so in an environment already used by this group, such as social networking sites. In this case, segmentation was not only beneficial for identifying more specific target audiences, but also for determining where a potential intervention might be effective.

APPLYING ANALYSIS TECHNIQUES

Data analysis doesn’t have to be complicated. Most descriptive statistics, and even some inferential statistics, can be completed in [Microsoft Excel](#), Google Sheets, or even with a calculator. More complex analyses require statistical software such as [IBM® SPSS®](#), [SAS](#), [Stata](#), or [R](#). Details on the underlying theory and step-by-step instructions on how to perform the various techniques described in this chapter are available in most statistics textbooks, as well as training activities and materials. Some useful references and resources for further learning are listed in Table 2.

Table 2: Textbook and online resources for further reading

Textbooks

Pallant, J. (2020). SPSS survival manual: A step by step guide to data analysis using IBM SPSS. Routledge.	An easy-to-read textbook on running data analysis using the statistical software SPSS
Tabachnick, B. G. and L. S. Fidell (2019). Using multivariate statistics. Boston, Pearson.	A key reference for running more complex multivariate statistical analyses
Field, A. Miles, J. & Field, Z. (2012). Discovering statistics using R. Sage Publications.	A guide to using the statistical package R for exploring and analysing data (sister publications are also available for SPSS and SAS software)
Online resources	
How2stats http://www/how2stats.net/p/home.html	An online guide with videos demonstrating step-by-step instructions for running a range of descriptive and inferential statistics in SPSS
Laerd Statistics https://statistics.laerd.com/	Online guides and tutorials for conducting statistical analyses in SPSS (free and paid service)
Institute for Digital Research and Education Statistical Consulting https://stats.idre.ucla.edu/	Resources for running more advanced statistical analyses in a range of software packages

Visualising data

Data analysis is often accompanied by tables and graphs such as bar charts and pie charts. Visualising data is helpful for quickly conveying key findings or messages. It is often easier to see patterns, trends, and outliers when numbers have been transformed into figures. While we can generate a correlation coefficient statistic to tell us the strength of a relationship between two variables, it can be easier to see this relationship when it's displayed on a scatterplot - e.g. if there is a strong positive relationship you would see the dots fall in a relative straight line from the bottom left corner to the top right; if there was no relationship you would expect to see virtually no pattern in the dots. Visualising the data would also help you see if a strong correlation coefficient statistic was being driven by only a few data points while the rest of the data did not have a very strong relationship.

Different data visualisation tools convey different sorts of information and are therefore appropriate for different types of data. For example, pie charts, boxplots, and bar charts are helpful for displaying descriptive statistics, whereas scatterplots and line graphs are better for depicting relationships between variables and are often used to accompany inferential statistics. See Figure 4 for an outline of the strengths and limitations of each of these tools.

If you're interested in data visualisation, keep an eye out for future BWA training opportunities on this topic.

	Tables	Pie charts	Boxplots	Bar charts	Line graphs	Scatterplots
Strengths	All data are provided in one place that users can access easily	Useful for depicting a part-to-whole relationship	Help summarise descriptive statistics (median, quartiles, outliers)	Bars can be easily and clearly compared side by side	Useful for depicting trends and/or predictions over time, where small changes in the slope of the line are easy to see	Useful for depicting relationships between continuous variables
Limitations	Can be too detailed and cluttered Main findings in the data are not immediately obvious	Difficult to read when slices are similar Not suitable when there are many slices, categories are non-mutually exclusive, or do not sum together	Do not show relationships between variables	Less useful for depicting the strength of the relationship between variables	Too many lines can be difficult to read Not suitable or potentially misleading when applied to some types of data (e.g., categorical)	Cannot summarise descriptive statistics Not appropriate for some types of data (e.g., categorical)

Figure 4: Strengths and limitations of common data visualisation tools used to explore behaviour change problems

Ask an expert

While data analysis, including audience segmentation, consists of an assortment of simple and complex techniques, involving a behavioural expert with quantitative training can be beneficial for several reasons, including:

- Searching the literature and consulting widely with experts in relevant fields to determine candidate criteria and strategies for segmenting;
- Designing and carrying out data collection, as well as preparing and cleaning the data for analysis;
- Carrying out complex analyses such as regression, cluster analysis, or more advanced techniques; and
- Ensuring that analytical approaches are robust and valid, that assumptions and other requirements are met, and that claims drawn from the data are appropriate and substantiated.

For example, as mentioned previously, we were engaged to conduct a suite of exploratory research in order to understand digital inclusion in Australia. This included undertaking a rapid review of academic literature on digital inclusion at a global level (Borg, Boulet, Smith, & Bragge, 2019), surveying a large representative sample of Australians to find out what digital inclusion looked like at a national level (Borg & Smith, 2017), and then conducting segmentation using cluster analysis to determine which Australians were at risk of being digitally excluded (Borg & Smith, 2018).

CONCLUSION

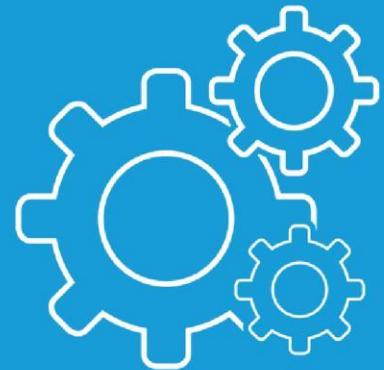
Data analysis is a powerful set of techniques that can help build a clearer picture of your specific problem. An evidence review (Chapter 1) can provide a landscape of prior knowledge, and systems thinking (Chapter 2) can draw on diverse perspectives and set up a framework for understanding complex problems. Adding to this breadth of knowledge, data analysis can provide a rich depth of knowledge about what is going on in a specific context and answer questions that have not yet been examined by other researchers or practitioners - providing a complementary approach for understanding the problem in the Exploration phase of The BWA Method.

Returning to our illegal dumping case study discussed throughout this chapter, Sarah learned a lot about her problem by exploring the data and applying a variety of data analysis techniques. Using descriptive statistics, she learned that the behaviour contributing most to her problem was illegal dumping of waste outside the charity stores. She then used inferential statistics (modelling) to identify the factors associated with illegal dumping at charity stores. Alternatively, Sarah could have segmented different charity stores using contextual characteristics – such as grouping stores based on accessibility (e.g. proximity to the road/footpath) or size (e.g. small, medium, large). She could have then explored the differences between the types of stores to get a better sense of where more and less waste was being dumped. Taken together, this not only helped to understand the problem (the Exploration phase of the BWA Method) but provided insights into the drivers of problematic behaviours (Deep Dive) and directed Sarah to targets for intervention options to reduce the problem behaviours (Application).

REFERENCES

- Alkiş, N. & Temizel, T. T. (2015). The impact of individual differences on influence strategies. *Personality and Individual Differences*, 87, 147–152.
- Australian Bureau of Statistics (October, 2020) *Statistical Language - What are Data?* ABS Website, accessed 18 February 2021.
- Borg, K., Boulet, M., Smith, L., & Bragge, P. (2019). Digital inclusion & health communication: A rapid review of literature. *Health communication*, 34(11), 1320-1328.
- Borg, K. & Smith, L. (2018). Digital inclusion and online behaviour: Five typologies of Australian internet users. *Behaviour & Information Technology*, 37(4), 367–380.
- Borg, K., & Smith, L. (2017). *Digital Inclusion: Report of Online Behaviours in Australia 2016*. Australia Post.
- Grunig, J. E. (1989). Publics, audiences and market segments: Segmentation principles for campaigns. In C. Salmon (Ed.), *Information campaigns: Balancing social values and social change* (pp. 199–228). Newbury Park: Sage.

- Haslam, N. (2019). Unicorns, snarks, and personality types: A review of the first 102 taxometric studies of personality. *Australian Journal of Psychology*, 71(1), 39–49.
- Wright, B. Smith, L. & Tull, F. (2018). Predictors of illegal dumping at charitable collection points. *Waste Management*, 75, 30-36.



THE METHOD

CHAPTER 4: Stakeholder consultation to improve behaviour change

By Dr Breanna Wright and A/Prof Peter Bragge

www.behaviourworksaustralia.org



INTRODUCTION - YOU HAVE TO TALK TO PEOPLE TO UNDERSTAND THEM

To help improve patient safety in Victorian public hospitals, the Victorian Managed Insurance Authority (VMIA) saw an opportunity to look at how regional and metropolitan hospitals were collaborating.

Patients in rural or regional hospitals sometimes need care that is not available at their local hospital, and they need to be transferred to another hospital to receive specialised care. This often involves transfers to a larger metropolitan hospital that can offer more services - for example, a CT scanner which is not viable in settings that don't have high patient volumes.

Delays in transferring patients can negatively impact patient care and patient outcomes. There are also technical and logistical challenges - for example, finding an ambulance that can transport a patient when there may be only one or two servicing a regional community.

However, to even set these parts of a patient transfer in motion, there first has to be a telephone call or conversation between two people - a doctor in a regional hospital with a sick patient in front of them, and another doctor in a metropolitan hospital dealing with other sick patients and limited beds. For the transfer to happen, both doctors have to agree on the need for the transfer and activate a plan.

Picture yourself as one of these doctors. What would be the things going through your mind during this telephone call? How might your other responsibilities be affected by the outcome of the call? How does the busy emergency department environment affect the conversation? What if you are really tired and nearly ending your shift - does this make the conversation different to when you've just started work?

These sorts of questions can't be answered by looking at protocols or guidelines that outline what should be done under certain medical circumstances. Understanding how such guidelines play out in the real world requires talking to the people involved, and listening to their experiences.

This chapter outlines the value of conversations such as this, which fit under the broad banner of *stakeholder consultation*. Specifically, we will explore the many ways in which stakeholder consultation can be undertaken to help understand and ultimately address behaviour change challenges, and outline some considerations for matching behavioural challenges to stakeholder consultation techniques.

STAKEHOLDER CONSULTATION THROUGHOUT THE METHOD

The BehaviourWorks Australia (BWA) Method, developed over ten years of applied behaviour change projects, represents an array of techniques that collectively address behaviour change challenges. The Method is not a "recipe" for every type of challenge - the techniques are used in various combinations according to which are most relevant for each individual problem. These techniques are presented across three overarching domains – "Exploration" to define the problem; "Deep Dive" to understand behaviours pertaining to the problem; and "Application" to test a behaviour change strategy based on the information gathered.

Stakeholder consultation can, and should, be utilised throughout all three of these domains - but particularly in the Exploration and Deep Dive sections.

In the Exploration phase, stakeholder consultation is a critical component of “unpacking the problem”.

Stakeholders bring their own tacit knowledge and understanding of the problem to the table.

“Stakeholder consultation” is a broad term covering a spectrum of consultation with groups relevant to the problem (Kvam, 2017). It is used to capture the views and perceptions of people who may be impacted by or have an interest in a project, and provides a way to take their views into account. Consultative processes with greater transparency and involvement of stakeholders can enhance trust, project acceptance, and local ownership (Kvam, 2017). It is worth noting that while we use the term “consult” throughout this chapter, this is used to broadly refer to all methods of consultation.

The aim of stakeholder consultation is to explore the multiple perspectives that help us build a picture of the problem and its context. Most problems will affect stakeholders differently and it’s important to understand these differences. This enables examination of a series of specific behaviours relating to the problem. By doing this, we can identify opportunities to influence the behaviours that can have positive impact.

Stakeholder consultation can also be used in the Deep Dive phase to unpack the drivers and barriers to specific behaviours or to identify appropriate intervention options.

The rest of this Chapter will predominantly discuss stakeholder consultation in relation to the Exploration phase of the Method.

WHY CONSULT WITH PEOPLE?

When key groups or perspectives aren’t included in the development of policies or programs, the best laid plans can go awry. In some cases, key groups are simply not consulted, leading to an inadequate picture of the problem. In 2014, Apple launched its “comprehensive” health tracker app. The app allowed people to track nearly everything; daily movement, exercise, heart rate, sodium intake, even copper intake! For such a comprehensive app, how did they forget to include a key health outcome for half their users - a period tracker? They didn’t have enough (or arguably any) input from women. Half of their potential users didn’t have a voice in what was important or useful to them in a health tracking app.

When it comes to solving difficult or complex health, social, environmental and organisational challenges, it’s important to know what the problem actually is. This may sound obvious, but as the Apple example shows, the truth is that many organisations rush to solutions without first conducting a detailed examination of the problem, which can lead to poor program outcomes.

Another key purpose of consultation is to challenge assumptions. Often researchers, policy makers and program designers have their own ideas about what will work. But these ideas come with their own assumptions. Assumptions about what end-users need or want, assumptions about how they will respond to different programs/policies, and assumptions

about what will have the biggest impacts on their lives. Without the insights of those whose behaviour can impact upon a problem, our own assumptions tend to fill the void.

It's easy to make assumptions about what people want or need if they are not there to correct you. For example, many assume that people who are paralysed following a serious spinal cord injury (SCI) want to be able to walk again. This makes intuitive sense, but is based on assumptions about the preferences of people with SCI. A survey of 681 people with spinal cord injury revealed different priorities. Those with quadriplegia (SCI affecting arms and legs) ranked arm and hand function first, with walking fifth out of seven of the listed priorities. For those with paraplegia (SCI affecting legs) the highest priority was sexual function, with walking ranked fourth out of seven (Anderson, 2004).

Consultation is also helpful in understanding how promising research findings from other parts of the world may translate into local settings. For example, there may be good evidence that a program that encourages people to get screened for breast cancer has worked in other countries, but how can we know if it will work here? What is the local context and considerations that we need to understand before knowing if something that has worked in other places, can work here too? To understand what will work to address problems or needs in a particular group or locale, consultation is imperative. Whilst other forms of data collection (see Chapter 3 of the Method) can also provide insights, talking to people, particularly those affected by the issue, often reveals invaluable knowledge about the local context, including the drivers and barriers of behaviour for the target population.

Importantly, consultation is also a critical factor to generating support for action. Involving people in all stages of the BWA Method - from understanding and defining the problem and examining behavioural drivers through to discussing options to address it - can generate support and engagement in the eventual solutions because people have been consulted from the beginning. Conversely, when people are made aware of interventions only when they are being implemented, they may be less receptive to the changes.

As these examples show, engaging with people directly affected or involved when addressing a behaviour change challenge can reveal insights and perspectives that can't be captured otherwise, including by existing knowledge sources such as academic publications or reports. It is only through proper stakeholder consultation that we can gain the insights needed to properly inform decisions, policies and programs. And avoid potentially costly assumptions.

IDENTIFYING THE PERSPECTIVES YOU NEED

Despite the importance and value of stakeholder consultation, it is not always done well. For example, it may be considered in an ad hoc manner, or viewed as a token part of the research development process.

Drawing upon previous parts of The Method, in particular Chapter 2: Systems Thinking and Behaviour Change,

will often help identify groups of stakeholders that should be engaged. This will likely involve a mix of **professionals** - those who have expertise or qualifications in the topic at hand - for example, doctors who transport patients between hospitals; and **citizens** - those with

experience or perspectives of the issue - for example, patients being transferred between hospitals and their families.

It's also important to capture the full spectrum of perspectives on the problem. This involves thinking broadly across factors like socio-demographics, but also how people are affected by the problem - for example, with different severity of injury or differing health needs requiring transfer to another hospital.

Mapping out the different perspectives on the topic or problem can help to make sure that diversity of opinion is captured. It's important to bear in mind that groups are not homogenous. So when considering groups of stakeholders, also give thought to the diversity within each group.

For example, in the introduction we considered the perspectives of two doctors. However, there are numerous other stakeholders that could also enhance our understanding of the situation and its associated behavioural implications. These include the nurses and other professionals that provide care; the patient and their family; the drivers of patient transfer vehicles; and hospital administrators who determine bed capacity and other operational matters.

TYPES OF STAKEHOLDER CONSULTATION

There are multiple ways to consult stakeholders. The chosen method should be tailored depending on the type of stakeholder. For the purpose of this Chapter, we will focus on the two groups defined previously - professionals and citizens.

The reason for considering these two groups separately is that citizens and professionals bring different, but complementary, perspectives to the consultative processes.

Professionals can engage with levels of detail relevant to their qualifications - for example, details of injuries and illness requiring hospital transfers, and research literature with specialist medical language. However, professionals lack direct experience of being an injured patient. Citizens can describe this experience and aid understanding of what it is like to go through a transfer from one hospital to another through their stories. However, they are less able to engage in specialist language and research. Both groups have valuable information to bring to consultation but neither have all the information needed to fully understand the challenge. They are therefore each in a position to contribute unique perspectives and enhance their understanding of the perspectives of others.

While consultation can be done with a combined group of different stakeholders, it's important to think about the aims of the consultation and if these can be achieved for each type of stakeholder within the group.

For example, we've found that including a few citizens as representatives in a group of professionals is not effective in really hearing from citizens, as the environment isn't conducive to citizens feeling able to speak up and share their experiences and opinions.

Placing consumer representatives (for example, from a peak body representing people with a particular condition) in a group of professionals partly addresses this problem. However, because they are representing a large group of people with diverse needs and perspectives,

there is less emphasis on individual stories. We have therefore found it more beneficial to hold separate stakeholder consultation sessions with citizens, followed by professionals sessions where consumer representatives are still included.

Engaging with citizens

There are multiple reasons and methods to engage citizens, consumers, or “end-users” (Eitzel et al., 2017). These can be conceptualized as a ladder from “light touch” forms of consultation to more in-depth approaches (see Table 1).

Table 1: Levels of consultation with citizens

IAP2'S PUBLIC PARTICIPATION SPECTRUM



The IAP2 Federation has developed the Spectrum to help groups define the public's role in any public participation process. The IAP2 Spectrum is quickly becoming an international standard.

PUBLIC PARTICIPATION GOAL	INCREASING IMPACT ON THE DECISION				
	INFORM	CONSULT	INVOLVE	COLLABORATE	EMPOWER
PROMISE TO THE PUBLIC	To provide the public with balanced and objective information to assist them in understanding the problem, alternatives, opportunities and/or solutions.	To obtain public feedback on analysis, alternatives and/or decisions.	To work directly with the public throughout the process to ensure that public concerns and aspirations are consistently understood and considered.	To partner with the public in each aspect of the decision including the development of alternatives and the identification of the preferred solution.	To place final decision making in the hands of the public.
	We will keep you informed.	We will keep you informed, listen to and acknowledge concerns and aspirations, and provide feedback on how public input influenced the decision.	We will work with you to ensure that your concerns and aspirations are directly reflected in the alternatives developed and provide feedback on how public input influenced the decision.	We will look to you for advice and innovation in formulating solutions and incorporate your advice and recommendations into the decisions to the maximum extent possible.	We will implement what you decide.

© IAP2 International Federation 2014. All rights reserved.

Source: International Association for Public Participation www.iap2.org.

Depending on the purpose or outcome needed from the consultation, different methods are appropriate. Table 2 provides a brief overview of common citizen consultation methods. Proper consideration should be given to why the consultation is being undertaken and what data or information from the consultations is most needed.

Table 2: Methods for citizen consultation

Method	Structure	No. of participants	Relevant outcomes
Individual approaches			

One-on-one interviews	Structured, semi-structured, unstructured Centred around specific research questions or topic	1	In-depth understanding of individual perspectives
Dyadic interview	Two participants interact in response to open-ended research questions	2	Useful to explore related/joint viewpoints (e.g. a patient and their carer)
Group approaches			
Focus groups	Interactive discussion	8-15 people	Non-sensitive topics where there are likely a range of viewpoints
Citizen panels	Facilitated discussion informed by a summary of relevant evidence	Approximately 15 people with varied lived experience of the problem	Engaging with evidence. Extracting values and preferences. Understanding the diversity of perspectives.
Citizen juries	Deliberations on evidence, followed by a voting system	20-100 people; stratified	Identifying priorities
Concept mapping	A series of maps are created to represent links between ideas and suggestions; maps are then refined	10-20 people	Decision-making, seeking consensus
Consensus panels	Provided with limited information on specific scenarios. Deliberations include rationale for decision	8-10 people	Seeking consensus
Nominal group techniques	Ideas are generated and then ranked amongst the group	Varied	Identifying priorities, seeking consensus

When selecting an appropriate method for citizen consultation, it's important to give due consideration to the purpose and outcome. By understanding the outcome that needs to be achieved, the right method can be selected and the consultation structured in a way that is able to elicit the desired information. Thus, it can be effective to start with the desired end result and work backwards.

Engaging with professional stakeholders: The Forum Approach

When it comes to consulting with professional stakeholders, BWA uses an overarching model known as the Forum Approach. It generally includes four activities: a rapid review, practice interviews, a citizen panel, and a structured stakeholder dialogue. Collectively, these activities promote the use of evidence to inform and drive policy and practice (ElJardali et al, 2014; Boyko et al, 2014).

As noted previously, a common shortcoming of policy and program development is the implementation of a "solution" without a detailed examination of the behavioural drivers that

need to be harnessed to optimise the chances of meaningful and sustained behaviour change. The Forum Approach overcomes this by combining research evidence, implementation insights, and consumer contributions with intervention design.

BWA used the Forum Approach as part of a three-year applied behaviour change research partnership with the Victorian Managed Insurance Authority (VMIA:

<https://www.behaviourworksaustralia.org/victorian-managedinsurance-authority/>). This program explored how simple, scalable interventions could reduce avoidable risks in delivering hospital care. One of the topics that was prioritised was “Improving hospital collaboration”.

As mentioned at the start of this Chapter, in many areas of Victoria, clinical care for patients with complex health issues is not available in small hospitals, as they are not designed to cater for all conditions (for example, they do not have the specialised equipment available in larger hospitals). Inter-hospital transfers therefore play an important role in the healthcare system, ensuring that patients can access appropriate care in a timely manner. Below is an example of how the consultation processes within the Forum Approach were applied to the challenge of how to improve the decision-making to transfer patients from a regional or rural hospital so that it is undertaken in a timely and efficient manner.

Rapid review of academic evidence

Rapid reviews provide a high level synthesis of relevant academic evidence (see Chapter 1).

In the hospital transfer project, the rapid review found that decisions to transfer patients are influenced by clinical factors, an increased chance of survival, quality of care and the need for a specific test or procedure. There is some asymmetry between how doctors and patients think about the decision to transfer; different factors are important to them and influence their decisions differently. There is a lack of guidance around how to navigate these transfers and so communication during transfer processes is not always optimal, creating barriers to effective communication and contributing to delayed transfers.

Practice interviews

Academic evidence is supplemented with qualitative interviews with experts who have deep knowledge of the context and problem. Engaging with people and practitioners “on the ground” can reveal invaluable information on how best-practice recommendations from research have played out in the real world and what barriers to implementation have been faced. This information allows us to respond not only to what the evidence says but to design a feasible and practical response that is tailored to local contextual factors.

In the hospital transfer project, we conducted interviews with two Ambulance Victoria staff, two Directors of

Medical Services, a consultant, a researcher, a Chief Medical Officer, a Quality and Risk Manager, a Patient Flow Coordinator, a Unit Manager, a Director of Clinical Operations and a Director of Nursing. Lack of skills, seniority and equipment were cited as contributing to delays in transfer decision-making. However, telemedicine, capability frameworks and protocols were reported to assist with timely decision-making. “Pushback” from metropolitan hospitals (for example, a reluctance to agree to the transfer due to heavy workloads or limited beds) was highlighted as another major factor.

Citizen Panels

Facilitated citizen panel sessions provide an in-depth understanding of the problem and its potential solutions from the viewpoint of citizens. Sessions are often developed in collaboration with a steering committee to ensure it meets the needs and requirements of all participants. The sessions aim to gain a shared understanding of the problem and its context informed by the evidence, practice and key contextual factors, culminating in an understanding of required behaviour change. A key benefit to citizen panels is not only connecting citizens to evidence, but also extracting underlying preferences and values that are important to citizens so that these can be incorporated into solution design.

In the hospital transfer project, we consulted with a panel of 15 Victorian community members with varied lived experiences related to transfers; including patients who had been transferred and their families/carers. Their main concerns centred around delays (due to bed availability, transport availability, indecisiveness, dismissal of patient and family concerns, or delays in diagnosis). They also had concerns about miscommunication between hospitals, such as test results not being available or the need to repeat tests.

Structured Stakeholder Dialogue

Structure stakeholder dialogues (Moat, et al, 2013) connect all the collated evidence from the previous stages (rapid review, practice interviews and citizen panel) with the people that can implement change. A structured stakeholder dialogue involves 12-18 experts, senior decision-makers and consumer representatives who are vested in the issue at hand. Prior to the dialogue, participants are sent the collated evidence. A half- to full-day facilitated discussion allows for deliberation on the issue and the evidence, and discussion of options to improve the issue. By inviting and encouraging stakeholders to lend their perspective and experience to the problem, informed and innovative solutions can be identified. Dialogues create space for “aha” moments to occur and yield a shared understanding of exactly who needs to do what differently and how that change might be achieved. Research has shown that this approach has very high participant satisfaction and crucially, leads to strong intentions to act on the basis of the evidence presented and discussed.

In the hospital transfer project, a day-long, structured stakeholder dialogue was convened. The dialogue was attended by 20 people representing government, policy, insurance, medicine, nursing, ambulance, operations, citizens and research. A briefing document summarising findings of a rapid review of academic evidence, practice interviews and citizen panel outcomes pertaining to this topic, was sent out to all dialogue participants in advance of the day.

The dialogue had three aims:

1. Gain a shared understanding of evidence, practice and key issues relating to interhospital transfer.

The dialogue found that the biggest challenges to inter-hospital transfers from regional and rural hospitals include:

- Lack of awareness of existing guidelines and protocols, and the role of hospitals within a network of other hospitals; lack of protocols and guidelines in some areas

- Lack of awareness about when to use various transfer options
- Lack of awareness and trust surrounding lower acuity hospital capabilities by higher acuity hospital staff
- Balancing the risks of managing a potentially deteriorating patient against the risks of transfer to a higher acuity facility (including opportunity costs of having ambulances out of local areas)
- Attitudinal barriers from potential receiving hospitals – for example, a culture of “my beds” and a default of “no” rather than “what plan can we put in place?”

2. Identify interventions to improve inter-hospital transfer from regional and rural hospitals that could be trialled and scaled across Victoria

The group deliberated upon a range of strategies to address identified behavioural challenges. Guidelines and protocols could be developed about where to send people and when to reduce the need for these discussions and improve decision making. The discussion unpacked how larger hospitals in each region could adopt a “buddy system” and take responsibility for patients from smaller hospitals within the region. Some regions reported doing this well and learnings could be drawn from these high-performing areas to improve behaviours in other regions. It was proposed that if larger hospitals could not take a patient, they should share ownership of finding another destination for the patient.

3. Prioritise an intervention and determine measures of success

A range of suggested options to improve inter-hospital transfers from regional and rural areas arose in the discussion. The options with the most support were:

- Trialling an awareness campaign to increase knowledge of transport resources, hospital capabilities and protocols.
- Trialling a communication strategy that reflects high-functioning systems to foster attitudinal change.

The Outcome

Based on the outcomes from this Forum Approach, we developed and tested the effectiveness of a video that reminded clinicians, who receive transfer requests, of the shared goals across hospitals and the healthcare system. The [persuasive video](#) contained a message to remind clinicians that they all want to help their patients get better.

After viewing the video, emergency department clinicians reported a stronger intention to accept a non-critical patient. They perceived the medical case as more severe and urgent, and were more accepting of the calling clinician’s report about their capability or resources to treat the patient.

Improving engagement between rural and metropolitan clinicians by creating a shared understanding and purpose can improve the efficiency of inter-hospital transport. Rural junior staff will be more comfortable asking advice when the metropolitan clinicians take a collaborative rather than a combative approach. – Tim Baker, Rural Emergency Physician

Lessons from successful Forums

The effectiveness of the Forum approach stems from a number of factors:

- All stakeholder dialogues are conducted under Chatham House Rule, which states that “participants are free to use the information received, but neither the identity nor the affiliation of the speaker(s), nor that of any other participant, may be revealed” (<https://www.chathamhouse.org/about-us/chatham-house-rule>) and no audio is recorded. This enables participants to provide input with the confidence that neither they nor their institution will be identified.
- There is no explicit mandate to reach consensus (although consensus can emerge from the discussion). This recognises that decision-making authority does not reside within the dialogue group, but within each of the organisations and groups that they represent. It also enables action to be taken despite the fact that not all individuals in a dialogue can commit to that action.
- High-level representation from all identified key groups builds a shared understanding of the complexities of the issue. Invitees are discouraged from sending junior proxies to the stakeholder dialogue, because high-level representation has more decision-making authority and influence in generating crucial buy-in for the next steps in addressing the issue.
- Connecting participants with the perspectives of others and relevant research evidence helps to challenge assumptions about what will work and generate broader thinking about possible solutions. Participants arrive at consultations with some ideas about what they want to contribute - this is part of the reason they have been invited in the first place - but it's incredibly useful to prompt new thinking by providing the evidence summary and by facilitating the sessions to ensure that all perspectives are heard. This is true both for stakeholders who are professionals and citizens.
- Observers at dialogues are strongly discouraged. Their presence can foster reservations in the participants and therefore provide a disincentive to provide a full and frank contribution to the discussion. We do not allow any observers in our dialogues for this reason - the only people in the room are the participants, facilitator and one other researcher who takes notes.

These features reflect key success factors to stakeholder consultation; free and open communication, realistic objectives for the discussion, having the right people in the room, and connecting people with evidence.

CONCLUSIONS

Stakeholder consultation is increasingly a critical component of research and solution design. Many funders now require consultation or co-design principles to be embedded in proposals. Along with utilising evidence, including from the academic literature (Chapter 1), understanding the system (Chapter 2), and other sources of data (Chapter 4), involving stakeholders from citizens to professionals enhances the relevance and integrity of research outcomes.

As we saw at the start of the Chapter, different stakeholders can have a very different understanding of a problem. Stakeholder consultation can not only help us gain an in-depth

understanding of those different perspectives but when people are brought together in a collaborative environment and hear not only from the evidence but also each other, it can help build a shared multifaceted understanding of the problem and can garner support for collective action.

REFERENCES

- Anderson, Kim D. 2004. "Targeting Recovery: Priorities of the Spinal Cord-Injured Population." *Journal of Neurotrauma* 21 (10): 1371–83.
<https://doi.org/10.1089/neu.2004.21.1371>
- Boyko JA, Lavis JN, Dobbins M. 2014. Deliberative dialogues as a strategy for system-level knowledge translation and exchange. *Healthcare Policy* 9 (4): 122.
- Broman, K.K., et al. 2018. Surgical Transfer Decision Making: How Regional Resources are Allocated in a Regional Transfer Network. *Joint Commission Journal on Quality and Patient Safety* 44 (1): 33-42.
- Craig, Simon S. 2005. Challenges in Arranging Interhospital Transfers from a Small Regional Hospital: An Observational Study. *Emergency Medicine Australasia* 17 (2): 124–31. <https://doi.org/10.1111/j.17426723.2005.00703.x>
- Eitzel MV, Cappadonna JL, Santos-Lang C, Duerr RE, Virapongse A, West SE, et al. 2017. Citizen science terminology matters: Exploring key terms. *Citizen Science: Theory and Practice* 2 (1).
- El-Jardali F, Lavis J, Moat K, Pantoja T, Ataya N. 2014. Capturing lessons learned from evidence-to-policy initiatives through structured reflection. *Health research policy and systems* 12 (1): 1-15.
- Kvam, R. (2017). Meaningful stakeholder consultation: IDB series on environmental and social risk and opportunity.
- Moat, Kaelan A, John N Lavis, Sarah J Clancy, Fadi El-Jardali, & Tomas Pantoja. 2013. Evidence Briefs and Deliberative Dialogues: Perceptions and Intentions to Act on What Was Learnt. *Bulletin of the World Health Organization* 92 (1): 20–28. <https://doi.org/10.2471/blt.12.116806>
- Ryan M, Scott DA, Reeves C, Bate A, Van Teijlingen ER, Russell EM, et al. 2001. Eliciting public preferences for healthcare: a systematic review of techniques. *Health technology assessment* 5 (5): 1-186.

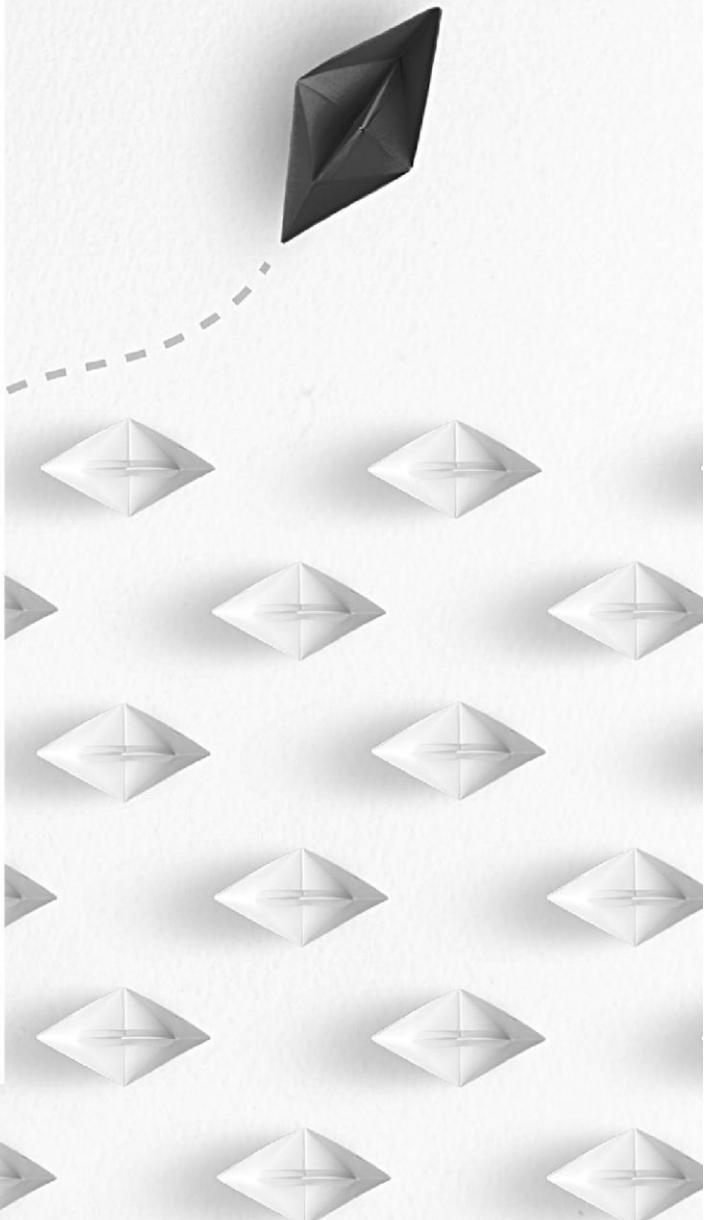
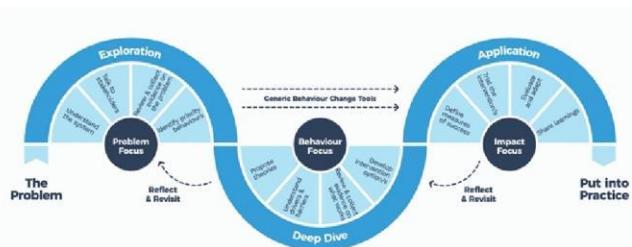


THE METHOD

CHAPTER 5: Getting ready to Deep Dive - defining, identifying and prioritising behaviours

By Sarah Kneebone, Mark Boulet, Lena Jungbluth, Jenni Downes and Celine Klemm

www.behaviourworksaustralia.org



INTRODUCTION

At the turn of the century, Australia was hit by severe drought. In the state of Victoria, a raft of legislative and infrastructure-development measures were introduced to maintain water supplies. Furthermore, water managers targeted householders in Melbourne through a series of high-profile campaigns to save water. But the list of different watersaving behaviours householders could perform seemed endless. These included behaviours to reduce personal water consumption (e.g., changing habits around toilet flushing, showering and clothes washing), to use water more efficiently (e.g., by installing water-efficient devices), or to diversify the water source (e.g., installing rainwater tanks). The question for water managers was “How do we select the ‘best’ behaviours to promote to householders to conserve water supplies?” Unfortunately, limited research existed at the time to inform their decision, which is why following the drought, a large-scale multi-disciplinary research program - the Cooperative Research Centre (CRC) for Water Sensitive Cities - was set up to identify the best approaches should such a crisis reoccur.

The BehaviourWorks Australia (BWA) component of this program, led by Liam Smith and Sarah Kneebone, identified 46 water saving behaviours for householders to reduce their water consumption and developed a tool to determine the “most promising” watersaving behaviours. To achieve this, they investigated existing water conservation literature, a series of Australian water-saving campaigns and engaged with experts from across the water sector (Kneebone et al., 2017). The approach the BWA team developed, which is now being used across a range of sectors and challenges, assists policy makers to answer the question “How can we prioritise a long list of potential behaviours to inform policy and practice?”

We know that many policy and program problems are complex, and do not have one single behavioural solution. Indeed, one problem could involve multiple different stakeholders, who may all need to enact a wide range of behaviours to create a positive impact. The previous chapters describe the Exploration phase of the BWA Method, including tools such as evidence reviews, systems mapping, working with data and engaging stakeholders to get a complete, thorough and holistic understanding of a particular problem. This chapter presents the “final” step of the Exploration phase, which is to **identify** and **prioritise** relevant behaviours that can have a positive impact on a problem. This critical step sets us up for the Deep Dive phase of the BWA Method, where we explore a priority behaviour from the perspective of a target audience and design an intervention to encourage the uptake of that behaviour.

This chapter will help practitioners identify the range of possible behavioural solutions to a complex problem and, just as importantly, select a priority behaviour to take through to the Deep Dive.

DEFINING BEHAVIOURS

Before thinking about identifying and prioritising behaviours, it is necessary to understand what behavioural scientists mean by the term “behaviour” and the criteria used to define a behaviour as specifically as possible.

A behaviour is an “**observable action**”, something we can see someone doing, and that is performed in a particular time and place (Fishbein and Ajzen, 2010). If we can see it, we can count it, and we can generate data (which is useful to find out if an intervention is working or not!). The key point is that when thinking about behaviour, the unit of analysis is what someone is *doing*, rather than what they might be thinking, believing or feeling. We are interested in the **action** rather than the internal state of mind of the individual.

It is also useful to focus on behaviours that are **desirable**. That is, what we want to happen rather than what we don't want to happen, again because it is easier to see (and count) when people are performing the correct behaviours rather than not performing the undesirable behaviour. For example, instead of “Don't take long showers”, we frame it as “Take four-minute showers”. We also focus on **end-state** behaviours rather than process behaviours. For example “Install a water-efficient dishwasher” rather than “Research the best water-efficient dishwasher for you”, “Go to the shop to buy a water-efficient dishwasher”, “Arrange delivery of a water-efficient dishwasher”.

Fishbein and Ajzen (2010) provided a framework of components for behaviour definition including the action, target, context and time. This was later modified to additionally specify the actor who performs the action (Presseau et al., 2019).

BWA uses an adapted version of this to help define behaviours:

WHO DOES WHAT with WHAT, WHERE and WHEN?

Each of these components makes up a different part of a behaviour. Table 1 outlines and defines these components, while Table 2 provides some examples of well-defined behaviours that have been identified to address specific problems.

Table 1: Components of behaviour

Fishbein and Ajzen (2010) and related versions	BWA version	Definition
ACTOR / AUDIENCE	WHO	The actor performing the behaviour. For example, the child, the mother, the politician. Who the audience is says a lot about how that action or behaviour is performed and to what extent it can change.
ACTION	DOES WHAT	The observable action that underpins the behaviour. For example, driving, swimming, walking or writing are all actions. They all involve a person physically doing something (that can be seen).
TARGET	WHAT	The target the action is directed towards. For example, the car when driving, or the pad and pen when writing, or the water when swimming. The target could be an object, person or anything towards which the action is performed.
CONTEXT	WHERE	The physical or environmental context or location in which the action is performed.

TIME	WHEN	The time when the action occurs. This could be very general (weekly or yearly) or much more specific (10 am each morning).
------	------	--

Table 2: Examples of well-defined behaviours

INDEX: WHO, DOES, WHAT, WHERE, WHEN

PROBLEM	BEHAVIOUR IDENTIFIED
Households consume large amounts of potable water, making water management in cities difficult during a drought.	Householders install a water-efficient dishwasher in their kitchen when renovating their house.
The office kitchen microwave is always filthy.	In the office kitchen, staff place a splatter guard over their plate when using the microwave to heat their lunch.
A large amount of food is thrown away by children at school.	Every day at school, children take their leftovers home at the end of the day so their parents can see what they have or have not eaten.
Melburnians are getting sick by swimming in Port Philip Bay when pollution levels are high after a rain event.	Swimmers check the EPA water quality app at home before going to the beach to see if the bay is safe for swimming.
The department has a culture of sexism and inequality.	Office manager directs male and female members of staff to take meeting minutes at the weekly staff meeting on an alternating basis, regardless of position every week.
Teenagers are not getting their full range of vaccinations to protect against cervical cancer.	Parents of teenagers submit permission slips on time when the school is running an immunisation program.
The rate of 000 calls for an ambulance in Victoria is increasing faster than the population is growing.	Victorians only call 000 for an ambulance in a time-critical, life-threatening situation.

Defining a behaviour using the AATCT (Audience, Action, Target, Context, Time) framework or the

WHO/WHAT/WHERE/WHEN criteria allows us to describe it as specifically as possible. We must be clear and precise about our behaviours for several reasons:

- Behaviours framed differently in terms of, for instance, their context or the time at which they occur will have different drivers and barriers to their adoption, which will impact on intervention design. For example, turning the bathroom tap off when brushing your teeth will have different drivers and barriers to turning the outside tap off when watering the lawn.

- Clarity around a target behaviour is important for the design and communication of behaviour change interventions, as our target audience must have an unambiguous idea of what we would like them to do. For example, asking Melbourne householders to save water is too generic; asking them to take fourminute showers provides a clear expectation and call to action.
- It is important to have a very clearly defined behavioural outcome measure for meaningful evaluation of an intervention. For example, we need to know whether a water conservation campaign based on installation of rainwater tanks has resulted in householders installing rainwater tanks.

Precise, detailed, and explicit descriptions of behaviour are critical, not only for understanding behaviour, but also for communicating with others what desirable or ideal behaviours actually look like. Clear behaviour definition is vital for successful project partnerships, research design, intervention development and program evaluation.

IDENTIFYING BEHAVIOURS

In 2017, the Western Australian Waste Wise Schools program identified a food waste problem in schools across the state. Several million whole fruits, sandwiches and other food items were being discarded, uneaten by students in primary and secondary schools. The Waste Wise Schools team wanted to design behaviour change interventions targeted at either students, teachers or parents that could reduce this problem.

When partnering with the project team, our first step was to conduct a rapid review of research evidence to identify the range of behaviours that have been associated with a reduction in food waste in schools. We found 12 different behaviours, some relevant to students, others to teachers and parents. These are shown in Table 3 (NOTE: to avoid overwhelming you with additional detail, we have not defined these behaviours as fully as described previously).

Table 3: Identified behaviours to reduce food waste

Audience	Identified behaviour
Parents	<ol style="list-style-type: none"> 1. Plan a week of school lunches for children 2. Pack school food based on children's hunger levels 3. Involve children in making their own lunches 4. Discuss with children what they like to eat and how much 5. Pack food that will keep and not spoil or go soggy 6. Include foods that can be eaten in stages 7. Purchase lunch boxes that reduce food damage and spoilage
Children	<ol style="list-style-type: none"> 8. Bring any uneaten food home from school

Teachers/schools	9. Implement a “take leftovers home” policy 10. Schedule eating time after playtime 11. Eat together as a class inside before playtime 12. Schedule longer lunch times
------------------	---

There are several insights to glean from this list that are important to this stage of identifying behaviours:

The list is made up of observable actions – you can “see” someone doing these actions, rather than being internal mental or emotional states. They also have an explicit ‘WHO’ and implied ‘WHERE’.

As already discussed previously, defining a behaviour as an observable action is foundational to a behavioural approach to any problem. It is especially important when it comes to monitoring the outcomes of behaviour change interventions, as we can more easily monitor changes in observable behaviours as opposed to changes in internal states.

The behaviours listed are all desirable – they are positive or solution oriented. Rather than saying “don’t put food waste in school bins”, we suggest that students could instead “take uneaten food home”.

Behaviours that are positive are often more informative than negative behaviours, which may not actually tell audiences what to do instead. Again, when it comes to monitoring the outcomes of behaviour change interventions, it can be easier to look for positive behaviours being performed, rather than the “absence” of negative behaviours.

Most of the behaviours listed are “end-state” or the “ultimate” behaviour that fixes the problem, rather than an early step, or chain of behaviours.

“Purchasing foods that will keep and not quickly spoil” is a prior step, but we are really interested in the end-state behaviour, which is whether they are packed for children to take to school, as this is what will reduce food waste in schools.

These last two aspects (making sure behaviours are desirable and end-state) become particularly important when moving to the next step of prioritising.

Approaches to identify behaviours

The Waste Wise Schools project used a **review of research literature and current practice** to identify potential behaviours that could be targeted. However, there are a number of other ways that practitioners can identify relevant behaviours:

- A recent [project](#) on community bushfire safety behaviours used **workshops with experts** from government fire management agencies to identify over 30 different behaviours.
- A project on prioritising householder circular economy behaviours reviewed **grey literature** such as government reports and campaigns to identify 99 initial “behaviours”, which were later condensed to 55.
- Our recent Waste Collaboration recycling [trials](#) used **system mapping** tools (see

Chapter 2) to identify multiple behaviours of three different groups of stakeholders.

Behaviour identification hints and tips

- You are looking for a long list! As long as the behaviours are observable, desirable and end-state, they should be on the list. Do not feel that you need to prioritise at this stage. If the behaviours are associated in some way with a positive influence on the problem, then include them.
- Be aware of the difference between desired project outcomes and the behaviours that need to be adopted to create those outcomes. Using the AATCT framework every time will help improve clarity.
- Behaviours can be tricky to “find”. While research and practice often broadly explores a problem, investigates a particular audience, or tests interventions, specific behaviours are often implicit and hidden (Kneebone et al., 2017). They may need to be filtered out of information in reports and research papers (look for outcome measures in surveys or trials) or generated by experts, who can be gently pushed to describe solutions as actions people can undertake.
- The behaviour identification process can take time. In our project on community bushfire safety, we ran two workshops with experts to identify behaviours and then refined the final list of behaviours with stakeholders over another two months.

PRIORITISING BEHAVIOURS

A long list of well-defined behaviours can provide a useful overview of what actions can be taken by different audiences to positively impact a particular problem. In any behaviour change project, it is important, however, to narrow down and select some key target behaviours to take forward into the next steps of the BWA Method.

Why do we need to prioritise?

The more potential target audiences and behaviours we can identify for a problem, the more relevant the question of prioritisation becomes. Depending on the problem at hand, the number of potential audiences and behaviours can vary greatly. At one end of the continuum, we can find simple problems which tend to have a small number of audiences and behaviours while at the other end we find wicked problems which tend to have multiple potential stakeholders (or audiences) and multiple potential behavioural solutions (see Figure 1).

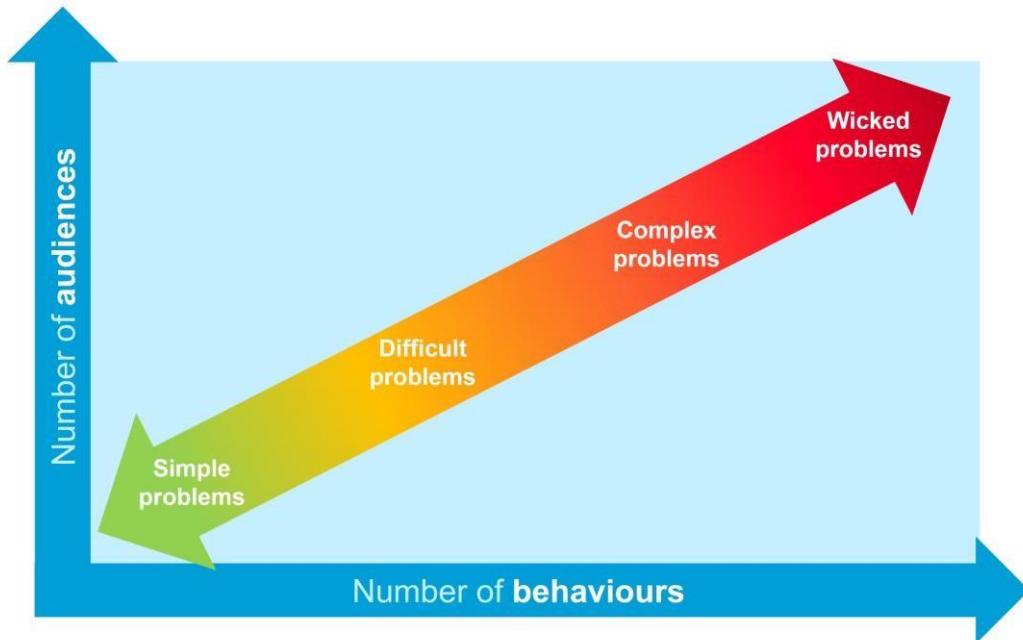


Figure 1: Number of audiences and behaviours for different problem types

Once a long list of relevant audiences and desirable behaviours to address a problem has been identified, it can be tempting to simply communicate these and expect behavioural adoption by the relevant groups. In the past, this “laundry list” approach has frequently been seen in the environmental space, where concerned citizens are exhorted to do “500 things to save the planet” or engage with “365 ways to save water”. This approach, however, has been criticised as confusing or disengaging, leaving the individual feeling overwhelmed. This can be at least partially explained by choice overload, where citizens or consumers are paralysed and unable to make a decision (Iyengar and Lepper, 2000), victims of the paradox of choice (Schwarz, 2004). The audience is left to engage with the behaviour they personally find the most straightforward or lowest effort - rather than those which will actually make a difference (Gardner and Stern, 2008).

Carrying a long list of behaviours into the Deep Dive section of the BWA Method risks resource and time constraints preventing a thorough behavioural diagnosis and thus impeding effective intervention design. Different behaviours may have different types or scale of impacts on the problem. For project efficacy it makes sense to pursue those with the greatest impact. Alternatively, particular behaviours might not be feasible for your specific program or agency to tackle, for various political or statutory limitations. Finally, some behaviours may have already been widely adopted by your target audience and further promotion of these would not achieve much additional impact.

For these reasons, it is important to narrow a long list of behaviours down to a smaller handful of priority actions that can be targeted for further research and intervention design. This does not mean the long-list is wasted! Indeed, it is a valuable resource that you can always come back to when those behaviours that were initially prioritised are being tackled through “live” interventions in the field.

Selecting priority behaviours

The BWA Method is set up as a series of questions to help guide project planning, with details of the tools and approaches needed to collect the evidence that informs the decisions

made around each question. The behaviour identification and prioritisation section of the Exploration phase is no different. Having identified a long list of behaviours, practitioners have to make decisions around which behaviours will be taken forward into the Deep Dive phase and behaviour diagnosis. The process requires four steps:

- 1) Selecting useful and relevant prioritisation criteria that are applicable for your problem and your project.
- 2) Collecting evidence or data on the criteria against your long list of behaviours.
- 3) Applying a prioritisation method to compare outcomes of your evidence collection.
- 4) Select a priority behaviour based on your criteria and evidence.

Step 1) Select your prioritisation criteria

There are many ways you can prioritise your long list of behaviours, from pragmatic decisions, to more involved, “analytical” processes. Which criteria you choose to inform your behaviour prioritisation should be determined by your own requirements or limitations, including what information you can collect, time and resource constraints and project expectations.

The most important thing is to apply your criteria to your identified behaviours to reach a point where you have just one or two final behaviours to take into the Deep Dive, or at least a ranked list that you can sequentially address.

BWA researchers typically use some of the following criteria to support decision-making around behaviour prioritisation:

- The **impact** of behaviours on the problem. This might include very specific measures of impact (such as how many litres of water might be saved in the home by performing a specific behaviour) to more general measures (such as a reduction in the risk of bushfires occurring or a reduction in the severity of their impact).
- The **ease** of performing a particular behaviour by the target audience, which influences how likely the audience is to adopt the behaviour. Measures of ease could include the physical effort, mental effort, or the time or financial cost the target audience faces to perform the behaviours.
- Current **adoption or participation** levels by the target audience of the behaviours. Depending on the problem, it may be useful to give a higher priority to those behaviours that are less commonly adopted.

Some of our partners have also included more pragmatic considerations when considering criteria for prioritisation. These might include the feasibility of targeting certain behaviours, how much they meet community expectation, whether they fall within an agency’s remit or their political sensitivity. For the food waste reduction behaviours listed previously, practitioners from the Waste Wise Schools prioritised not just on the perceived impacts of the behaviours but also on the weight of the evidence base behind each behaviour (Boulet et al., 2019).

Step 2) Collect data to inform prioritisation

Once you have selected the criteria that you will be using to prioritise the behaviours in your long list, the next question is how you will assign a “score” to each of the criteria. While the simplest way might be to give each behaviour a subjective or intuitive ranking based on your own experience, there are some other possibilities for you to consider:

- For a recent project that attempted to prioritise recycling and waste reduction behaviours, we utilised an **expert elicitation workshop** in which relevant program and policy officers assigned behaviours a score against each of the criteria used, and then discussed the scores and their reasoning, before giving a final rating. This process helps experts to take into account a broader range of evidence when rating.
- Our community bushfire safety project sent out a **survey to experts** across the region, asking them to rate a behaviour’s impact based on its highest possible impact, its lowest, and their final best guess. This process helps people arrive at a better intuitive estimate.
- A number of BWA studies have also used large-scale **community surveys** to establish adoption and likelihood scores by directly asking people whether they currently perform the behaviour, whether they have performed the behaviour in the past, and how easy (or hard) they think it would be to adopt the behaviour themselves, considering for example the physical or financial effort required. This process helps gain an understanding of the target audience’s perceptions around behaviour adoption.

Other sources of information include using research, grey literature, or publicly accessible databases (such as the Australian Bureau of Statistics or the Cochrane Database of Systematic Reviews) to collect data to help score different criteria such as impact on the issue or current adoption of a behaviour.

Step 3) Apply a prioritisation method

Once you have collected some data or evidence against your selected criteria, there are a number of formal tools that you can use to help your prioritisation process. Most prioritisation tools are a form of **Multi-criteria analysis**. These allow you to score each behaviour against specific criteria, for example, “1 to 5”, with 1 being a low score and 5 being a high score, or “high, medium, low”. Behaviours can then be ranked by total score, with the top scoring behaviours being most desirable as project priorities. For example, **Community Based Social Marketing** uses a cost-benefit analysis to refine long-lists of behaviours, scoring behaviours on their impact on the issue, probability of adoption by the target audience and existing participation to form a single, ranked list (McKenzie-Mohr, 2011).

Other tools might relate to specific settings or contexts. The **NERO** model was developed by Liam Smith and Jim Curtis as a result of investigating preferences of zoo visitors for participation in biodiversity conservation behaviours (Smith et al., 2010). This approach also combines scores to provide a behaviour ranking, based on four key criteria:

- Novelty of the behaviour: visitors preferred to be asked to do something new, such as recycling their old mobile phone to protect gorilla habitat.
- Ease of performing the behaviour: actions that were low effort for a visitor to engage with.
- Response efficacy: there was a direct line of sight between the behaviour the visitor was asked to perform and the impact on the problem (e.g., donations to feed an animal received better support than a donation to the zoo generally).
- On-site: visitors preferred behaviours in the here and now, things they were asked to do at the zoo which they could engage with during their visit.

These combined score approaches can be very helpful to rank a long list of behaviours. However, a single score outcome risks losing some of the nuance within the data. When making a final decision around behaviour selection, it can be important to understand how behaviours compare across the prioritisation criteria. Individual projects will determine whether behaviour impact or the likelihood of adoption (or other criteria) are more important for decision making.

BWA's **Impact Likelihood Matrix** is another prioritisation method, but produces a visual representation of how each behaviour is scored on the selection criteria (Kneebone et al., 2017). The Matrix can provide a useful summary of the evidence which is easily interpretable by decision makers and retains the nuance of multiple data sets. BWA has used the Impact Likelihood Matrix in a range of projects to assist behaviour selection in projects relating to recycling, food waste avoidance at home, water conservation in the home, research use by teachers and community bushfire safety behaviours. An example of the Impact Likelihood Matrix being applied to identify priority behaviours is given in the case study below.

Step 4) Select a priority behaviour

In the past, projects and research outputs have suggested prioritising target behaviours through identification of the “*low-hanging fruit*”; behaviours which have a high likelihood of participation (as they are seen to be easy) and have a high impact on addressing the problem (Inskeep and Attari, 2014). However, practitioners need to consider the impact of existing participation rates on these options. If the vast majority of your target audience is already performing the target behaviour, it may not be economic to promote it further. In addition, it can be important to be wary of high ease behaviours, with high potential likelihood of adoption but little impact on the issue. Thøgersen and Crompton (2009) caution against the promotion of such “*simple and painless*” behaviours. Any decisions should be informed by selection criteria relevant to the desired outcome and the problem at hand.

Prioritisation Tool case study: The Impact Likelihood

Matrix What is the Impact Likelihood Matrix?

The Impact Likelihood Matrix is a visualisation approach to behaviour prioritisation which involves mapping behaviours onto a simple graph by their prioritisation criteria scores. Behaviours are mapped by their scores on the **impact on the issue** (i.e., the extent to which

behaviour uptake would solve the problem) on the y-axis, and the **likelihood of behaviour adoption by the target audience** on the x-axis. Once all the behaviours have been mapped by their scores, the graph can be overlaid with a two-by-two matrix which places each behaviour within one of four quadrants. The location of each behaviour within the quadrants, and relative to other behaviours on the graph, is then used to support behaviour selection (see Figure 2).

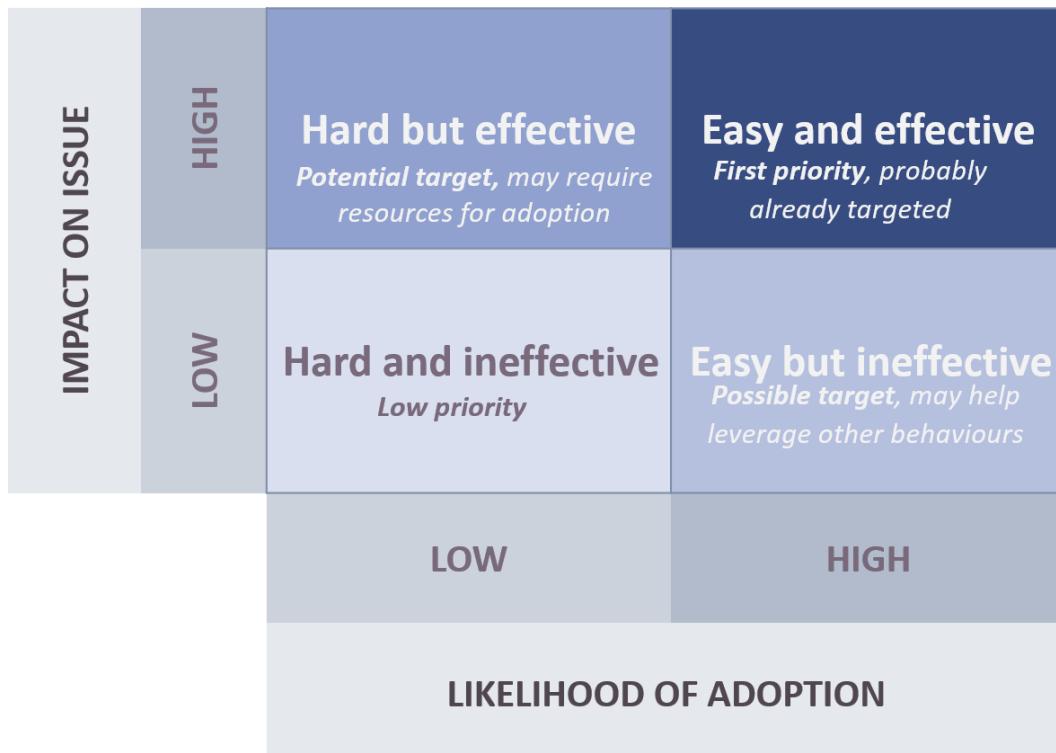


Figure 2: The four quadrants of the Impact Likelihood Matrix (Kneebone et al., 2017)

As Figure 2 shows, the quadrants illustrate how behaviours can then be prioritised:

- **Low** impact on the issue, **low** likelihood of adoption by the target audience: on the whole, behaviours in this quadrant will be a low priority. They don't contribute much to solving the problem and the target audience is less likely to engage in them.
- **Low** impact on the issue, **high** likelihood of adoption: this quadrant indicates a slightly higher priority. These behaviours again won't contribute a great deal to solving the problem, but your target audience is likely to take them up. It is worth bearing David MacKay's words here "If we all do a little, we will achieve a little, we need to achieve a lot" (MacKay, 2008).
- **High** impact, **low** likelihood: these behaviours could represent longer term goals, or areas that need investment in terms of quality intervention development, as they offer a high contribution to addressing your problem but work is needed to encourage their uptake by the target audience.
- **High** impact on the issue, **high** likelihood of adoption: these are the "low-hanging fruit" and potentially the highest priority as these behaviours make a high contribution to addressing your problem and your target audience is likely to engage in them.

However, you may need to consider existing participation rates. **Using the Impact Likelihood Matrix for prioritisation**

We used the Impact Likelihood Matrix to help unpack the 46 household water conservation behaviours identified from the literature and expert elicitation described in the Introduction.

Steps 1 and 2: We used multiple data sources to create the scores needed to populate the graph - see Table 4, including data from different audiences. The impact data was derived from water professionals' perceptions, while the likelihood of adoption data and current participation rates were drawn from householders themselves.

We also ran a comparison to investigate how much perceptions differed or were common to the two audiences (professionals vs householders) and found they were largely in agreement around impact on the issue, but varied around perceptions of ease. Practitioners should be mindful that experts' assessments of impact and likelihood of adoption and other prioritisation criteria may differ from those of the target audience. Input directly from the target audience may therefore be a useful sense-check.

Table 4: Summary of the different data types and evidence sources used to populate the Impact Likelihood Matrix for water saving behaviours

	Impact of behaviour adoption on the issue	Likelihood of behaviour adoption	Current participation
Data source	Online survey with water professionals	Online survey with householders	Online survey with householders
Measure used	Perception of behaviour impact on water saving	Perception of effort required to participate in the behaviour	Self-report of current participation in water saving behaviours

Step 3: We used the collected data to build a prioritisation graph with the Impact Likelihood Matrix overlaid, see Figure 3. In addition to impact on the issue and likelihood of adoption, we also included criteria around current participation. To reflect this, each behaviour is represented on the matrix with a circle, with the circle size indicating current behaviour adoption by the target audience. The larger the circle, the fewer people are engaged in the behaviour, and therefore the greater potential for change. Finally, we were interested in distinguishing behaviours by their type (using different colours in the matrix), because habitual behaviours are harder to change and therefore may be less easy to target as a project priority (Verplanken & Roy, 2016). These layered data provide detailed information for decision makers to base their selection on and demonstrate how multiple criteria can be applied for prioritisation.

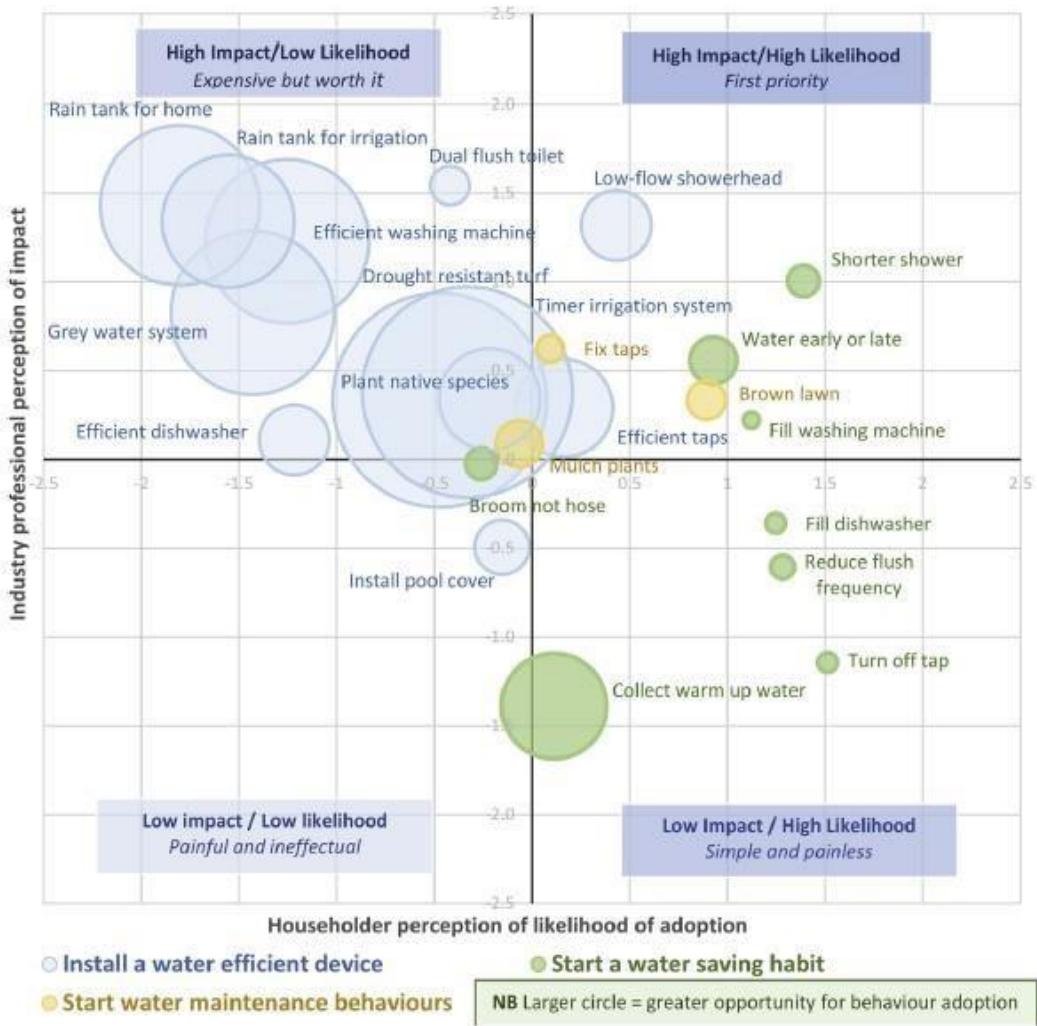


Figure 3: The Impact Likelihood Matrix for household water conservation behaviours, including current behaviour participation rates and behaviour type (Kneebone et al., 2017)

Step 4: For water managers considering how to develop policy and practices to mitigate the impacts of the next big drought to hit Australia, the results of the Impact Likelihood Matrix provide some clear direction for investment. The graph in Figure 3 suggests that most of the high impact/high likelihood “low hanging fruit” behaviours already have a high participation rate, and involve hard-to-shift habitual behaviours, such as *taking a shorter shower* or *filling the washing machine before use*. However, high impact/low likelihood, installation-type (and therefore one-off) behaviours such as *installing a water efficient washing machine* or *installing a rainwater tank for irrigation use*, or *installing a grey-water system*, have a much greater opportunity for change with lower current participation rates. Therefore, if urgent water saving measures in Australia are required in future, these behaviours represent promising candidates to focus water managers’ intervention efforts.

A final remark: Practitioners can use an Impact Likelihood Matrix to easily see which is the most promising behaviour from their long list and prioritise this to take forward into the Deep Dive Phase of the BWA Method for behaviour diagnosis (Kneebone et al., 2017). Importantly, this visual tool provides an easily interpretable justification for decision-making and rationale to provide to senior stakeholders. Although this example of the matrix uses impact and likelihood as the selection criteria, practitioners can build their own matrices using

their own selection criteria, scoring behaviours in terms of evidence and data available for their own context. The same method can be applied to provide an indication of potential behaviour preferences for a range of criteria, from financial cost to existing participation to scores for project relevance or department applicability. The matrix approach for behaviour selection is valuable due to its flexibility and relative simplicity.

CONCLUSION

The identification of a priority behaviour (or behaviours) is a critical step in the BWA Method for behaviour change project planning. Without a clearly defined target behaviour, projects will lack focus and clarity, and behaviour diagnosis as well as intervention design is rendered impossible. This chapter presented the final step of the Exploration phase of the Method. By applying the AATCT framework (or WHO DOES WHAT with WHAT, WHERE and WHEN) to clearly define a behaviour, practitioners can collate a long-list of potential behaviours to positively influence a particular problem. Relevant behaviours can be identified from a number of sources including expert knowledge, the literature or system maps. To support the decision process and narrow the longlist of behaviours down to a smaller number of target behaviours, practitioners can select from a number of criteria and prioritisation approaches.

Addressing the water managers' challenge mentioned in the Introduction, in this chapter, we used the example of BWA's Impact Likelihood Matrix (Kneebone et al., 2017) to illustrate how water managers can prioritise the most promising water saving behaviours to promote to householders during a drought from a long list of behaviours.

Once a priority target behaviour has been identified, this sets you up for the Deep Dive section of the BWA Method, in which we explore a target behaviour from the perspective of a target audience and design an intervention to encourage the uptake of that behaviour.

REFERENCES

- Boulet, M., Wright, B., Williams, C. & Rickinson, M. (2019). Return to sender: A behavioural approach to reducing food waste in schools. *Australasian Journal of Environmental Management*, 26(4), 328-346.
- Fishbein, M., & Ajzen, I. (2010). *Predicting and changing Behavior: The Reasoned Action Approach*. New York, NY: Psychology Press, Taylor & Francis Group.
- Gardner, G. T. & Stern, P. C. (2008). The short list: The most effective actions US households can take to curb climate change. *Environment: Science and Policy for Sustainable Development*, 50(5), 12-25.
- Inskeep, B. D., & Attari, S. Z. (2014). The water short list: The most effective actions US households can take to curb water use. *Environment: Science and Policy for Sustainable Development*, 56(4), 4-15.
- Iyengar, S. S. & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, 79(6), 995-1006.

Kneebone, S., Smith, L. & Fielding, K. (2017). The Impact-Likelihood Matrix: A policy tool for behaviour prioritisation. *Environmental Science & Policy* 70, 9-20.

MacKay, D. (2008). Every big helps. In *Sustainable Energy without the Hot Air* (pp.114-117). Accessed from www.withouthotair.com (last accessed 27 April 2021).

McKenzie-Mohr, D. (2011). *Fostering sustainable behavior: An introduction to communitybased social marketing* (3rd ed.). Gabriola Island, BC: New Society Publishers.

Presseau, J., McCleary, N., Lorencatto, F., Patey, A. M., Grimshaw, J. M., & Francis, J. J. (2019). Action, actor, context, target, time (AACTT): A framework for specifying behaviour. *Implementation Science*, 14, 102.

Schwartz, B. (2004). *The paradox of choice: Why more is less*. New York: Ecco.

Smith, L., Curtis, J. & Van Dijk, P. (2010). What the Zoo should ask: The visitor perspective on pro-wildlife behavior attributes. *Curator*, 53(3) 339-357.

Thøgersen, J. & Crompton, T. (2009). Simple and painless? The limitations of spillover in environmental campaigning. *Journal of Consumer Policy*, 32, 141-163.

Verplanken, B., & Roy, D. (2016). Empowering interventions to promote sustainable lifestyles: Testing the habit discontinuity hypothesis in a field experiment. *Journal of Environmental Psychology*, 45, 127-134.



THE METHOD

CHAPTER 6: "You are not normal!" Understanding the influences on behaviour

By Jim Curtis, Morgan Tear, Filia Garivaldis
and Fraser Tull

www.behaviourworksaustralia.org



INTRODUCTION

In the south-west region of Victoria, Australia, there are over 5,000 wetlands, many of which are on private farmland. With the drying climate, activities such as cropping (i.e., the planting, growing and harvesting of crops such as wheat, barley, canola) have moved into the region, posing a risk to wetlands, as farmers are confronted with declining profit margins and the need to maximise the use of their land (including planting crops on wetland areas). To address this challenge, the Glenelg Hopkins Catchment Management Authority (GHCMA) and other local community groups have attempted to engage farmers on the importance of protecting wetlands, often focusing on their rich biodiversity and environmental benefits (e.g., water purification, drought and soil erosion protection, reduction of greenhouse gases). But such efforts only reached 10-20% of the region's landholders, and the threat of cropping on local wetlands remained. In response, the GHCMA and other local community groups wanted to explore how the behavioural sciences could help with this challenge. They went through a process of identifying and prioritising behaviours, and landed on the desired behaviour of "farmers programming GPS coordinates into machinery to avoid wetlands when cropping." But to influence this behaviour, the GHCMA knew that their previous persuasive efforts needed to take a different approach. Instead of assuming what would influence farmers to change their behaviour, they needed to put these assumptions aside and ask farmers about what behavioural influences mattered to them. In other words, a "Deep Dive" was required.

Once we have identified and prioritised the behaviour(s) we want to change (based on the principles outlined in the previous chapter), we might be tempted to jump straight to interventions, only to discover that the interventions don't work as well as we would have hoped, or have had the unintentional effect of making the problem worse (Osman et al., 2020). Some examples of interventions that have created these unintended consequences include signs in Arizona's Petrified Forest National Park that increased the theft of petrified wood (Cialdini et al., 2006), the introduction of fines at a child care centre that increased rather than decreased late pick-ups (Gneezy & Rustichini, 2000), and a new campaign where zombies were used to encourage disaster preparedness behaviours that was less effective than previous messages (Fraustino & Ma, 2015). In the end, our interventions might have targeted the wrong influences on behaviour or made the wrong assumptions about what would matter to our target audience.

This chapter marks the beginning of the Deep Dive phase of the BWA Method. This phase focuses on understanding and collecting audience data on the different influences on behaviour, reviewing evidence of "what's worked" elsewhere to promote similar behaviours (applying the same principles outlined in Chapter 1), and using these collated insights to design interventions. Remarkably, this phase of understanding the audience that links the target behaviour with the choice of intervention is often overlooked, and so the choice of intervention strategy is instead based on implicit or assumed models of behaviour change (Michie, van Stralen, & West, 2011).

In this chapter, our focus is on introducing you to the different influences on behaviour. While this aligns most closely to the "propose theories" segment of the BWA Method, hinting at understanding different theories of behaviour, we have purposely decided not to go down a theory-heavy path in this chapter. Instead, and at the risk of being accused of being a bit self-indulgent, we thought we would introduce you to the different influences on behaviour

based on one of the chapter author's (Jim Curtis) own research journey in the behavioural science field (hence the switch from "we" to "I" in the next section of the chapter). Following this, we will then provide you with a couple of examples where the influences of behaviour have been synthesised into a finite set of variables that can be used to kick-off the Deep Dive phase.

WHERE DID IT ALL BEGIN?

"You are not normal. They're not like you!". These are the words of one of my PhD supervisors, and have played an important role in my behaviour change research journey. We have repeated these words in multiple training deliveries to students and practitioners, and have found that this simple collection of words has resonated with them just as much as they did with me all those many (too many!) years ago.

What do these words mean? Well, they are about not assuming that what would influence YOU to change behaviour would influence OTHERS the same way. For example, as a behaviour change researcher with somewhat "green" tendencies, it would be remiss of me to assume that farmers will respond to messages or appeals that focus on the rich biodiversity assets of wetlands the same way I would as a reason to protect them. While some might, others will consider other influences or factors that are more important to their decisionmaking and behaviours. Our job as researchers is to put ourselves in their shoes (via audience research) to understand what matters and influences them, and use these valuable insights to design interventions.

But putting aside our own assumptions can be a challenge. Most of us have our own intuitive understanding of behaviour and why people do the things they do. And for good reason. We are easily able to consider our own behaviour and then readily ascribe reasons for why we did it. Think about how you commuted to work this week, or what you ate for lunch yesterday, or a product you bought recently. If someone were to ask why you performed these behaviours, no doubt there is a raft of reasons you can articulate. However, assuming those same reasons apply to others can be dangerous, and might prove to be incorrect. Looking at behaviour through different lenses of known influences (based on a plethora of theories of behaviour from psychology and the social sciences) can therefore be a more objective way to understand why people behave the way they do. Ultimately, by embarking on a Deep Dive, it is about "increasing the odds" we are targeting the right behavioural influences relevant to our target audience (rather than our own).

So where did my behavioural science research journey begin? It was as a research assistant on a project at [Port](#)

[Campbell National Park](#) in Victoria (one of Victoria's most highly visited national parks because of the Twelve Apostles landform). The problem? Visitors walking off the designated track, causing erosion of the surrounding vegetation and landscape, and putting themselves at risk due to the steep and high ocean cliffs (Figure 1). While signs existed encouraging people to stay on the walking track (with messages warning people of unstable cliffs, snakes, or telling visitors to "show you care", with no further information about "what" to care about), evidence and observations of people walking off the track remained. Our task was to develop some new messages that could be displayed on temporary signs to see if they worked "better" to influence people to stay on the walking track.



Figure 1: Visitor walking off the track at Port Campbell National Park (source: Jim Curtis)

So, what did we do? We conducted interviews and surveys with visitors on-site (while also battling a fly plague at the time), and asked them about the positive/negative outcomes of staying on the track (attitudes), who would approve or disapprove of them staying on the track (subjective norm), and what factors or circumstances make it easy or difficult for them to stay on the track (perceived behaviour control). For those of you who are familiar with behavioural theories, you will probably recognise this project involved an application of the Theory of Planned Behaviour (Ajzen, 1991). From the audience research, we discovered certain beliefs that were common to “compliers” (i.e., those who stayed on the track) and “non-compliers”, while other beliefs were different between the two groups. These insights were then used to craft five temporary signs, which we implemented in the field. Survey research and observations took place on-site to see which signs worked the best in encouraging people to stay on the walking track (and influenced their underlying attitudes, norms, and perceptions of control, as well as engagement levels with the signs).

So, what did we find? Well, in the spirit of “you are not normal”, the leaders of the project (Professors Sam Ham and Betty Weiler - who would later become my PhD supervisors) asked the park managers in a workshop to make predictions on the results. They showed them the signs we tested, and asked those in attendance to participate in a blind vote. They were asked to nominate which sign they thought worked the best, and which sign they thought worked the worst in terms of influencing visitors’ underlying attitudes, norms and perceived behaviour control, as well as visitor intentions to stay on/walk off the track. The vote revealed that park managers expected that a sign that focused on telling visitors to “do the right thing” to protect the environment and featuring a picture of the local park ranger would work the best. In contrast, they expected the worst performing sign was one that focused on “don’t miss out on the best view or photo”, conveyed by a fictional professional photographer in a letter of thanks to park authorities for designing a great track to get the best photos. Turns out that the park managers had trouble thinking like a visitor - the field results revealed the opposite to what park managers expected. While both signs focused primarily on attitudes, one represented the attitude of park authorities, while the other focused on an attitude of visitors (where visitors had told us they went off the track in the hope of getting a better view or photo). And while there was creative licence in how we

presented the messages (informed by interpretation research - for example, using a personal anecdote), not to mention the ethical debate of using a fictional person on a sign (which continues to be a point of great discussion in our behaviour change training deliveries), I remember leaving that session knowing that “You are not normal. They’re not like you!” would become a mantra in my future behaviour change research efforts.

What happened next?

After this experience, I was hooked on the behavioural sciences and wanted to learn more (and combine it with my passion for the outdoors). So, I was fortunate enough to get a PhD scholarship from the former Cooperative Research Centre for Sustainable Tourism in Australia, and worked with two park agencies to try and influence visitors to use alternative transportation systems (in this case, shuttle buses) when visiting national parks. Like any PhD, I had to do a literature review on existing research and theory (with a particular focus on visitor and park management). The focus of the review was also influenced by my leaning towards using the Theory of Planned Behaviour (TPB), given my experience from Port Campbell National Park. I soon realised that one of the core strengths of the TPB - that is, it represents a succinct and tested framework of variables that impact on people’s intentions and behaviour - could also be seen as one of its shortcomings in terms of not considering other potential influences. Many researchers have therefore tested possible extensions of the theory (e.g., by adding measures like self-identity, anticipated regret, habit) to see if they could do a better job of explaining behaviour (which the original author of the TPB - Icek Ajzen - always acknowledged as a possible approach in applications of the theory).

So, I was keen to expand my understanding of other behavioural influences (both during and after my PhD), to see how other researchers have approached and investigated this task. If I was to nominate some phases of understanding where I had some “pivotal moments” of “enlightenment” when it comes to understanding behaviour (beyond those articulated in the TPB), I would summarise these under the following headings:

Norms are not just about the approval of others

One of the earliest books I read during my PhD was Robert Cialdini’s “Influence” (Cialdini, 2009). A multitude of new editions have been released since, but I still think this is one of the most engaging books I have read on behaviour change, and highlights that the influence of “norms” is not just about getting the approval of others, but what others are doing “like us” in similar situations, as well as how we want to be perceived by others. His principles of persuasion (Box 1) offer a highly accessible toolbox of behaviour change considerations, which can often be applied when we don’t always have the opportunity to do a Deep Dive. And I would also argue that the use of norms (based on the research of Cialdini and many of his co-authors and collaborators - e.g., Noah Goldstein, Steve J. Martin, Jessica Nolan, Wesley Shultz, among others) remain one of the most pervasive and widespread behaviour change techniques that we witness and experience every day in our society. From billboards and bills to online and in-store sales techniques, norms are everywhere (prompting recent

debates that ask the question whether we might be becoming inoculated to the influence of norm-based messaging).³

BOX 1: ROBERT CIALDINI'S SIX PRINCIPLES OF PERSUASION

1. RECIPROCITY: People feel obliged to give back to others in the form of a behaviour, gift, or service that they have received first.
2. SCARCITY: People want more of those things they can have less of.
3. AUTHORITY: People follow the lead of credible, knowledgeable experts.
4. CONSISTENCY: People like to be consistent with the things they have previously said or done.
5. LIKING: People prefer to say "yes" to those that they like.
6. CONSENSUS: When uncertain, people will look to the behaviours of others to determine their own.

Breaking (bad) habits

I would then nominate the role of "**habits**" as my next key learning moment (and I'm stealing some content from my co-author here - Fraser Tull - who specialises in habits). When people do a behaviour often enough, and in the same context, they can end up performing the behaviour mindlessly without any real thought or deliberation. For example, most of us don't have to think about the pros and cons of brushing our teeth at night. Instead, it's just something we do before heading into bed. As you can imagine, this makes breaking "bad" habitual behaviour difficult, as people may slip into performing the behaviour before they've even realised they're doing it. This poses problems for many behaviour change interventions, as they're often based on the assumption that people have complete control over their behaviour. When habits are involved, we often see an initial improvement in response to an intervention, followed by a quick return to baseline levels. There are, however, things we can do to increase our odds of success. Interventions have shown to be considerably more effective when implemented during periods of natural disruption. Educational and incentive programs, for example, are more effective when people are moving house or moving their place of work (Verplanken & Roy, 2016). During these times, the environmental triggers are gone, and so habit is no longer driving the behaviour. Interventions that aim to establish positive alternative habits, while also tackling bad habits, have also been shown to have lasting effects. For example, obese patients are more likely to maintain their weight loss when provided with strategies based on habit theory (e.g., have

³ According to Osman et al. (2020) and their review of behaviour change interventions that fail, the most common intervention used that resulted in failures were social norming or social comparisons. However, they acknowledge it is difficult to ascertain whether this failure rate is simply a case that social comparisons and social norms don't always work (e.g., because not all social norms are created equally and are known to backfire if used inappropriately), or that they are the most commonly used behavioural interventions (because they are cheap; they can be applied across multiple contexts), so this increases the likelihood of failures being reported.

an apple at the beginning of your lunch break), compared to standard treatment information (e.g., eat more fruit) (Gardner et al., 2021).

BOX 2: DID YOU KNOW?

In research by one of the world's leading authors on habits - Bas Verplanken - interventions promoting environmentally sustainable behaviours were more effective among recently relocated residents to the city of Peterborough in the UK compared to existing residents. The window of opportunity for these interventions to be effective was three months after relocation, after which the disruptive influence of relocation was no longer present (Verplanken & Roy, 2016).

And how long does it take to form new habits? Well, it depends on the behaviour. Research by Lally, van Jaarsveld, Potts, and Wardle (2010) revealed that a state of automaticity from repeatedly performing behaviours was reached between 18 and 254 days depending on the behaviour, with a median of 66 days. So, when you want to form new habits, make sure your behaviour interventions are running long enough for these new habits to become established.

The emotion of emotions

When BWA was established in 2011, one of our founding partners was (and remains) [The Shannon Company](#) - a specialist behaviour change company where the focus is on delivering campaigns to corporate, government and community clients for environmental and social good.

A key part of their work is using **emotions** to [inspire](#) (and sometimes [confront](#)) people to willingly change their behaviour. Seeing how these campaigns impact people in terms of their own emotional response has been something we have witnessed time and time again when examples of these campaigns are shown in presentations or in training. Up until that time, I probably hadn't truly appreciated or witnessed how influential emotions can be. This could partly be attributed to the fact that emotions are not predictable in how they directly or indirectly influence behaviour (either positively or negatively). While there are different "camps" of people who advocate whether a positive emotional frame (e.g., joy, hope) or a negative emotional frame (fear, anger, anxiety) should be used to instigate change, the direction of response to these different emotions remains unpredictable (Chapman, Lickel, & Markowitz, 2017). Emotions can also interfere with other potential influences on behaviour, and we often test this with the following exercise in training. Imagine you have been asked by a local wildlife sanctuary to pick up roadkill (dead animals that have been hit by vehicles) the next time you see some when driving (acknowledging that local road authorities and emergency service personal might rightly warn against performing such a behaviour). We don't reveal "why" the sanctuary is asking this (that comes later). We then ask participants what could be some of the outcomes (attitude) of doing this behaviour. Typical positive outcomes include rescuing a baby animal that might be in its dead mother's pouch (if the animal is, say, a kangaroo) and preventing vehicle collisions if a driver of a car swerves to avoid the dead animal. Typical negative outcomes include putting themselves in danger when moving the animal, delaying their journey, and whether they might later become ill from handling the animal. We then ask how they might "feel" about picking up roadkill, and that's when responses related to disgust about seeing, moving and smelling roadkill come to

the fore. Our final question is “Would you pick it up?”, and while some people are still up for the challenge, others realise their negative emotional response would get in the way of any previous positive attitudes. This interplay between emotions and attitudes is certainly not unique to this pairing of influences, but serves as a reminder that emotions are part of the overall behaviour change equation - you just need to be mindful that their direction of influence can be positive and negative, and will vary across behaviours and contexts.

And so why did the local wildlife sanctuary want people to pick-up roadkill? To stop endangered [wedge-tailed eagles](#) eating the roadkill and getting killed by cars in the process.

Making us conscious of the unconscious

Around the same time that BWA was established, two books were released by leading behavioural scientists. The first was “Nudge” by Richard Thaler and Cass Sunstein (2008), and the other was “Thinking, Fast and Slow” by Daniel Kahneman (2011). Both pointed to two different systems of thinking when it comes to decision-making and behaviour - one characterised by fast, instinctive, and automatic thinking, while the other is slower, deliberative and calculating. On a daily basis, we spend most of our time making decisions and behaving without much conscious thought (e.g., travel mode choice, food choice, household chores, brushing our teeth - although when it comes to the latter, some students I have delivered training to admit that brushing their teeth is not an unconscious decision and will weigh up whether to skip the morning ritual that day. They're not like me!!).

Unfortunately, a lot of attempts at behaviour change tend to assume people are paying attention and are logical in their decision-making, and ignore the unconscious and more automatic elements of human decision-making.

What authors such as Thaler, Sunstein and Kahneman did (along with others such as Dan Ariely, Angela Duckworth and Katy Milkman, to name just a few) was to bring to life the cognitive biases (systematic and predictable patterns of deviation from assumed rational judgements) based on decades of prior research, and harness our knowledge of the workings of these biases to influence people to make better decisions. Terminology such as loss aversion, sunk cost fallacy, confirmation bias, hyperbolic discounting and anchoring (among many, many others) suddenly became a more present component of our behaviour change toolbox, as such biases are often not collected when conducting audience research, as individuals are typically not conscious of these unconscious influences.

BOX 3: WOULD YOU TAKE THIS BET?

You are offered a gamble on the toss of a coin

If the coin shows tails, you lose \$100 If

the coin shows heads, you win

\$150 Would you take this bet?

Most people's answer to this question represents a demonstration of "loss aversion".

With cognitive biases, sometimes we just need to know which ones might be impacting on our priority behaviour. For a project where we wanted to influence people in financial stress to seek advice early from a financial counsellor, we identified the following biases that might be at play (in a negative way):

- OPTIMISM BIAS ("Everything will work out")
- SUNK COST BIAS ("I've invested so much already that I can't stop now")
- STATUS QUO BIAS ("I don't want to change the way I live")
- THE BANDWAGON EFFECT ("Going into debt is normal")
- LOSS AVERSION ("Making sacrifices today feels so much bigger than equivalent savings in the future")
- CONFIRMATION BIAS ("Everything I read about financial counselling confirms my own negative beliefs")
- PRESENT BIAS ("Smaller pay-offs today are better than larger ones in the future")
- OSTRICH EFFECT ("I don't want to hear or read anything about me being in financial stress")

Don't underestimate context

When it comes to the influence of context on behaviour, we see it as covering two realms: (1) choice architecture - the structures and architecture in the environment that influence behaviour (and our habits); and (2) culture - the shared social expectations that exist beyond any individual or single behaviour. Both of these realms include a mixture of physical, social, temporal and symbolic/cultural features that can influence a person's behaviour in direct or indirect ways. A fly sticker on a urinal (physical aspect) can ensure that men take better aim when urinating, the fact that we tend to wear similar clothes to work as those that we work with (social aspect), and the considerable increase in charitable donations during Christmas (cultural aspect) are all examples of how these realms can influence behaviour.

It can be useful to think about context through the lenses of proximity and distance.

Proximate contexts describe those that have a more direct, immediate effect upon behaviour - the physical location and social environment where a behaviour is taking place, like the workplace, the school, the supermarket, or the home. Features of context that are more distant – the unemployment rate, a region's climate, a news cycle extending for months on a particular topic – can influence behaviour more indirectly, and are usually more difficult to

observe. We need to also be aware that what defines proximity and distance is not so much based on physical distance, but on its salience to the individual. For example, we tend to prioritise things that are “here and now” over things that are further down the track. On another farmer project where we were exploring the drivers and barriers of using organic compost, a subsequent drop in the milk price meant a shift in focus among farmers (that is, just trying to earn a living rather than the “nice to haves” of considering alternative compost choices). While the milk price change for farmers may have been the result of things occurring away from their farms (e.g., global downturn in milk prices, oversupply, and the bargaining power of major supermarkets), it became a very proximate influencer based on its salience to the farmers.

In our descriptions of behavioural influences so far, many focus on the individual. However, some of the work we do focuses on influences at the organisational level - for example, how does the culture of an organisation impact its ability to deliver priority outcomes; what discrepancies exist between espoused organisational values and enacted values; and how can we implement organisational change interventions that avoid adding to the frightening statistic that most organisational change programs simply don’t work? I have yet to come to any clear answers on these, but I continue to learn from my colleagues who specialise in this space (including my coauthor of this chapter Morgan Tear) and believe this is a behaviour change frontier that will remain a priority, with the Royal Commission into Misconduct in the Banking, Superannuation and Financial Services Industry a stark reminder of this space (indeed, in many of our training deliveries, participants regularly bring an organisational change problem they want to solve).

BOX 4: ACCELERATORS OF ORGANISATIONAL CHANGE

One of my earliest forays into the organisational change world involved reading the work of John P. Kotter. While there are numerous organisational change texts out there, I still find his eight steps to accelerate change in an organisation a useful grounding:

1. Create a sense of urgency around a single big opportunity.
2. Build and maintain a guiding coalition.
3. Formulate a strategic vision and develop change initiatives designed to capitalise on the big opportunity.
4. Communicate the vision and the strategy to create buy-in and attract a growing volunteer army.
5. Accelerate movement toward the vision and the opportunity by ensuring that the network removes barriers.
6. Celebrate visible, significant short-term wins.
7. Never let up. Keep learning from experience. Don’t declare victory too soon.
8. Institutionalize strategic changes in the culture (Kotter, 2012).

So, these have been some examples of my learnings on the influences of behaviour SO FAR. And I say “so far”, because the learning never really stops, as both researchers and

practitioners continue to develop, test and discover new things about people and behaviour across different contexts, and increase the body of evidence (both successes and failures) that we can draw from to design more effective behaviour change interventions (Osman et al., 2020).

SYNTHESES OF BEHAVIOUR INFLUENCES

While “we” (switching back from the “I” focus of the previous section) have purposely steered clear of providing individual accounts of different behavioural theories and their pros and cons, we do see value in drawing attention to syntheses of behaviour influences that describe a finite set of variables that could be used in any behavioural diagnosis or Deep Dive. Such syntheses draw on decades of prior research and evidence that have been devoted to developing and testing theories of behaviour.⁴ One of the earlier attempts was undertaken by Fishbein, Triandis, Kanfer, Becker, and Middlestadt (2000). These leading theorists agreed that for a person to perform a given behaviour, one or more of the following variables must hold true:

1. There is a strong positive **intention** (or commitment) to perform the behaviour.
2. There are no **environmental constraints** that make it impossible for the behaviour to occur.
3. The necessary **skills** are present to perform the behaviour.
4. A positive **attitude** towards performing the behaviour exists (i.e., the advantages or benefits of the behaviour outweigh the disadvantages or costs).
5. There is more perceived **social pressure** to perform the behaviour than to not perform the behaviour.
6. The performance of the behaviour is perceived as consistent with a person’s **self-image**.
7. There is an overall positive **emotional** reaction to performing the behaviour.
8. The person perceives that he or she has the **capacity** (i.e., perceived self-efficacy) to perform the behaviour under a number of different circumstances.

Later attempts at this process have essentially mirrored the above variables, with some slight variations. For example, the Theoretical Domains Framework summarises 33 psychological theories, and identifies 14 domains or influences on behaviour (Cane, O’Connor, & Michie, 2012). The framework aims to make behaviour change theory more accessible to researchers and practitioners (Table 1).

⁴ Theories of behaviour provide an integrated summary of constructs, procedures and methods for understanding behaviour, and present an explicit account of the hypothesised relationships or causal pathways that influence action (Michie & Abraham, 2004). While not being exhaustive in terms of accounting for a full range of possible determinants (they are designed to be deliberately simple), they provide behaviour change researchers and practitioners with a means of avoiding implicit assumptions when selecting appropriate intervention strategies. Some take a more generalist approach, offering a concise account of universal constructs that can be applied to most behavioural domains. Examples include the Theory of Planned Behaviour (Ajzen, 1991) and its successor the Reasoned Action Approach (Fishbein & Ajzen, 2010). In contrast, others focus on particular behavioural determinants or behavioural domains, such as the Theory of Self Efficacy (Bandura, 1977) and the Value-Belief-Norm Theory of Environmentalism (Stern, Dietz, Abel, Guagnano, & Kalof, 1999).

Table 1: Theoretical Domains Framework (Cane et al., 2012)

DOMAIN	DESCRIPTION
KNOWLEDGE	An awareness of the existence of something
SKILLS	An ability or proficiency acquired through practice
SOCIAL/PROFESSIONAL ROLE OR IDENTITY	A coherent set of behaviours and displayed personal qualities of an individual in a social or work setting
BELIEFS ABOUT CAPABILITIES	Acceptance of the truth, reality, or validity about an ability, talent, or facility that a person can put to constructive use
OPTIMISM	The confidence that things will happen for the best or that desired goals will be attained
BELIEFS ABOUT CONSEQUENCES	Acceptance of the truth, reality, or validity about outcomes of a behaviour in a given situation
REINFORCEMENT	Increasing the probability of a response by arranging a dependent relationship, or contingency, between the response and a given stimulus
INTENTIONS	A conscious decision to perform a behaviour or a resolve to act in a certain way
GOALS	Mental representations of outcomes or end states that an individual wants to achieve
MEMORY, ATTENTION AND DECISION PROCESSES	The ability to retain information, focus selectively on aspects of the environment and choose between two or more alternatives
ENVIRONMENTAL CONTEXT AND RESOURCES	Any circumstance of a person's situation or environment that discourages or encourages the development of skills and abilities, independence, social competence, and adaptive behaviour
SOCIAL INFLUENCES	Those interpersonal processes that can cause individuals to change their thoughts, feelings, or behaviours
EMOTION	A complex reaction pattern, involving experiential, behavioural, and physiological elements, by which the individual attempts to deal with a personally significant matter or event
BEHAVIOURAL REGULATION	Anything aimed at managing or changing objectively observed or measured actions

At BWA, in both our research and training, we tend to draw on a synthesis of behavioural influences by Andrew Darnton and the UK's Government Social Research Service (Darnton, 2008). They identified over 60 models of behaviour, and identified a range of common behavioural variables among these models. We have adapted these slightly and are summarised in Table 2.

Table 2: Influences on behaviour, adapted from Darnton (2008)

INFLUENCE	DESCRIPTION
ATTITUDES	Overall favourable or unfavourable evaluation of engaging in a behaviour
SOCIAL NORMS	Social rules that indicate what are the common, expected and acceptable behaviours in a particular situation.
CAPABILITY	Personal physical, financial or psychological capabilities to undertake the behaviour
OPPORTUNITY	Factors or circumstances beyond the individual that provide them with the means to carry out the behaviour
HABITS	Repeatedly performed behaviours in stable contexts with little thought or deliberation
EMOTION	Actual or anticipated feelings in response to performing a behaviour
BIASES	Systematic and unconscious tendencies to think, decide and behave in certain ways, leading to predictable deviations away from a perceived standard of rationality or good judgement
CONTEXT	Structures and architecture in the environment that influence behaviour
	Shared social and cultural expectations that exist beyond any individual or single behaviour

Influences on behaviour can also be classified along various dimensions, ranging from conscious to unconscious, malleable or fixed, person-specific or situation specific.

Influences can also be classified in accordance with their strength of impact on behaviour. That is, some influences are remote, where they impact behaviour at an abstract level, such as personality traits and values (such influences are often classified as “background factors” in theoretical frameworks, and therefore manifest in more proximal influences, such as attitudes, norms etc.). At BWA, we often classify the influences in Table 2 across conscious and unconscious dimensions. Attitudes, social norms (at certain times), capability and opportunity tend to exert an influence on behaviour that we are more conscious of. The other four - context, biases, emotion, and habits - tend to exert an influence on behaviour that we are less conscious of (we also argue that social norms tend to have both conscious and unconscious influences on our behaviour). The extent to which any of these influences are relevant depends on who is performing the behaviour, what the action is, and where and when the behaviour is taking place (see Chapter 5 for how to define a behaviour).

TIPS FOR PRACTITIONERS

Generating an appropriate level of understanding or insight into some or ideally all of these influences will increase the chances of a behaviour change practitioner (or researcher) identifying the key influences related to a priority behaviour, which can later be targeted in an intervention. However, the relative influence of these determinants will ultimately depend on the behaviour. For example, behaviours that are perceived as easy and convenient (but

are not being performed) are potentially more likely to be influenced by factors that sit outside capacity and opportunity considerations, while behaviours performed in public and are subject to visible levels of scrutiny are arguably more susceptible to the influence of social norms. We also can't assume that the relative impact of these influences will be the same for everyone in our target audience. It might therefore be worthwhile segmenting our audience (see Chapter 3) based on different behavioural influences, and design interventions that respond to these differences.

Ultimately, a behaviour will be influenced by more than one factor, and so focusing a behaviour change campaign based on addressing just one factor will inevitably underestimate the range of potential influences on the behaviour. Understanding these different influences can therefore assist practitioners to effectively identify the suite of factors that need to be considered when developing an intervention program.

While theories of behaviour do not specify how to change behaviour, the choice of behaviour change technique or intervention should nevertheless be based on a core understanding of the most significant influences that determine the target behaviour. However, the reality is that many intervention designers pay insufficient attention to analysing the nature of the behaviour as a starting point for a behaviour change intervention (Michie et al., 2011). In the absence of this knowledge, Stern (2011) explains that the choice of behaviour change technique or strategy is relegated to implicit (assumed) theories of change - that people can be counted on to follow regulations and norms, that they will do what is economically most advantageous, and that new and efficient technologies are readily adopted. Under these circumstances, the choice of intervention is often based on assumptions that these theories of change adequately capture behavioural reality. While likely to contain some elements of truth, none are nearly complete, and are likely subject to behavioural "myths" that will lead to intervention failure.

WHAT DID WE FIND OUT FROM THE FARMERS?

Returning to our opening case study, we conducted a series of in-depth interviews with farmers in south-west Victoria, and asked them some broader questions around how cropping and wetland protection could co-exist on private land, as well as the potential influences on the behaviour of using GPS coordinates to program farm machinery to avoid wetlands when cropping. We found at the outset that there was a lack of clarity on what constitutes a "wetland", so the prospect of protecting something that was somewhat undefined or unclear presented an initial **capability** challenge. We then explored whether farmers considered wetlands to be an asset or liability on their farms in terms of delivering positive or negative outcomes (i.e., **attitudes**). Frequently mentioned positive outcomes included wetlands providing grazing opportunities for livestock (especially during hot or drought conditions), being run-off areas during heavy rains that protect the crops from being flooded, and offering biodiversity rich habitats for beneficial insects, birds and other wildlife. However, there were negative outcomes as well, such as wetlands providing habitats for pests (e.g., foxes, feral cats, weeds) and detracting from the farm value and profitable acres of the farm (if unable to crop), especially among the increasingly tight profit margins faced by farmers (**context**). Such contrasting views also manifested in different decisions among farmers to crop (or drain) wetlands or protect them. Those that cropped saw efficiency and profit gains from using all their available land, while those that didn't saw the opportunity to save wasting time, inputs and resources from trying to grow a crop in a wetland area (in

addition to the benefits expressed previously). Other influences included farmers feeling a level of distrust towards government authorities (fearing they will take away their rights to their land), while trusted sources tended to be other farmers and agronomists (**social norms**). Many also conveyed a sense of pride (**emotion**) as being custodians of the land, with many farms being owned by the same family for multiple generations.

When we asked farmers specifically about using GPS coordinates to program farm machinery, all agreed it was an easy thing to do (**capability** and **opportunity**, as there were no obstacles in terms of skills and availability of equipment to do this), and potentially had the added benefits of reducing driver fatigue by not requiring continual manual adjustments and/or allowing inexperienced staff to better operate machinery. However, beliefs around inefficiencies and lost opportunities to generate a viable return on their land remained.

In addition to the more conscious and deliberative insights we gathered from the interviews (and it was clear that farmers often engage in a highly calculated decision-making process), we also explored some potential unconscious influences (**biases**) that had been documented elsewhere in the literature when it comes to engaging farmers. We found examples of programs where personalising communication and connecting it to the local environment was an important precursor for engaging farmers, as well as using farmers (rather than government authorities) as credible sources of information (while also conveying a sense of “social proof” that other farmers are adopting the desired behaviours). The strong connections or networks between farmers have also seen the idea of using group incentives (i.e., where incentives are only offered based on an agreed level of threshold participation) being used to engage farmers, with the incentives being withdrawn if certain targets are not met (drawing on **loss aversion** and **social norms**). We also discovered opportunities for the **emotion** of “anticipated regret” - for example, painting a future picture where wetlands could provide income earning opportunities through government carbon offset schemes, only to realise (regret) that these opportunities had been lost by failing to protect wetlands in the present.

As you can see, farmers’ decisions to crop or protect wetlands was influenced by a complex mixture of behavioural influences, well beyond previous engagement efforts that had adopted an altruistic approach to argue for the protection of wetlands. Understanding and targeting such influences will therefore be imperative in redesigning future engagement and behaviour change efforts.

CONCLUSION

By the end of reading this chapter, we hope you might realise the truth that “You’re not normal”. That is, don’t assume that what might influence you to change your behaviour will be the same for others. Indeed, embrace the view that those you are trying to influence are the “normal” ones, and so your task is to understand what matters and influences them when it comes to undertaking your prioritised behaviour.

While this chapter has introduced you to some recognised influences on behaviour by offering you examples of syntheses describing these influences, the question remains as to how you go about collecting data on these influences. That question (among others) will be answered in the next chapter.

REFERENCES

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179-211.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191-215.
- Cane, J., O'Connor, D., & Michie, S. (2012). Validation of the theoretical domains framework for use in behaviour change and implementation research. *Implementation Science*, 7(1), 37.
- Chapman, D. A., Lickel, B., & Markowitz, E. M. (2017). Reassessing emotion in climate change communication. *Nature Climate Change*, 7(12), 850-852.
- Cialdini, R. B. (2009). *Influence: Science and practice* (5th ed.). Boston: Allyn and Bacon.
- Cialdini, R. B., Demaine, L. J., Sagarin, B. J., Barrett, D. W., Rhoads, K., & Winter, P. L. (2006). Managing social norms for persuasive impact. *Social Influence*, 1(1), 3-15.
- Darnton, A. (2008). *GSR behaviour change knowledge review. Reference Report: An overview of behaviour change models and their uses*. London: HMT Publishing Unit.
- Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach*. New York: Psychology Press.
- Fishbein, M., Triandis, H. C., Kanfer, F. H., Becker, M., & Middlestadt, S. E. (2000). Factors influencing behavior and behavior change. In A. Baum, J. E. Singer, & T. A. Revenson (Eds.), (pp. 1-17). Mahwah: Lawrence Erlbaum Associates.
- Fraustino, J. D., & Ma, L. (2015). CDC's use of social media and humor in a risk campaign—"Preparedness 101: Zombie Apocalypse". *Journal of Applied Communication Research*, 43(2), 222-241.
- Gardner, B., Richards, R., Lally, P., Rebar, A., Thwaite, T., & Beeken, R. J. (2021). Breaking habits or breaking habitual behaviours? Old habits as a neglected factor in weight loss maintenance. *Appetite*, 162, 105183.
- Gneezy, U., & Rustichini, A. (2000). A fine is a price. *The Journal of Legal Studies*, 29(1), 117.
- Kahneman, D. (2011). *Thinking, fast and slow* (1st ed.). New York: Farrar, Straus and Giroux.
- Kotter, J. P. (2012). Accelerate! *Harv Bus Rev*, 90(11), 44-52, 54-48, 149.
- Lally, P., van Jaarsveld, C. H. M., Potts, H. W. W., & Wardle, J. (2010). How are habits formed: Modelling habit formation in the real world. *European Journal of Social Psychology*, 40(6), 998-1009.
- Michie, S., & Abraham, C. (2004). Interventions to change health behaviours: Evidencebased or evidenceinspired? *Psychology and Health*, 19(1), 29-49.

- Michie, S., van Stralen, M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, 6(1), 42.
- Osman, M., McLachlan, S., Fenton, N., Neil, M., Lofstedt, R., & Meder, B. (2020). Learning from Behavioural Changes That Fail. *Trends in Cognitive Sciences*, 24(12), 969-980.
- Stern, P. C. (2011). Contributions of psychology to limiting climate change. *American Psychologist*, 66(4), 303.
- Stern, P. C., Dietz, T., Abel, T., Guagnano, G. A., & Kalof, L. (1999). A value-belief-norm theory of support for social movements: The case of environmentalism. *Human Ecology Review*, 6(2), 81-97.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven: Yale University Press.
- Verplanken, B., & Roy, D. (2016). Empowering interventions to promote sustainable lifestyles: Testing the habit discontinuity hypothesis in a field experiment. *Journal of Environmental Psychology*, 45, 127-134.

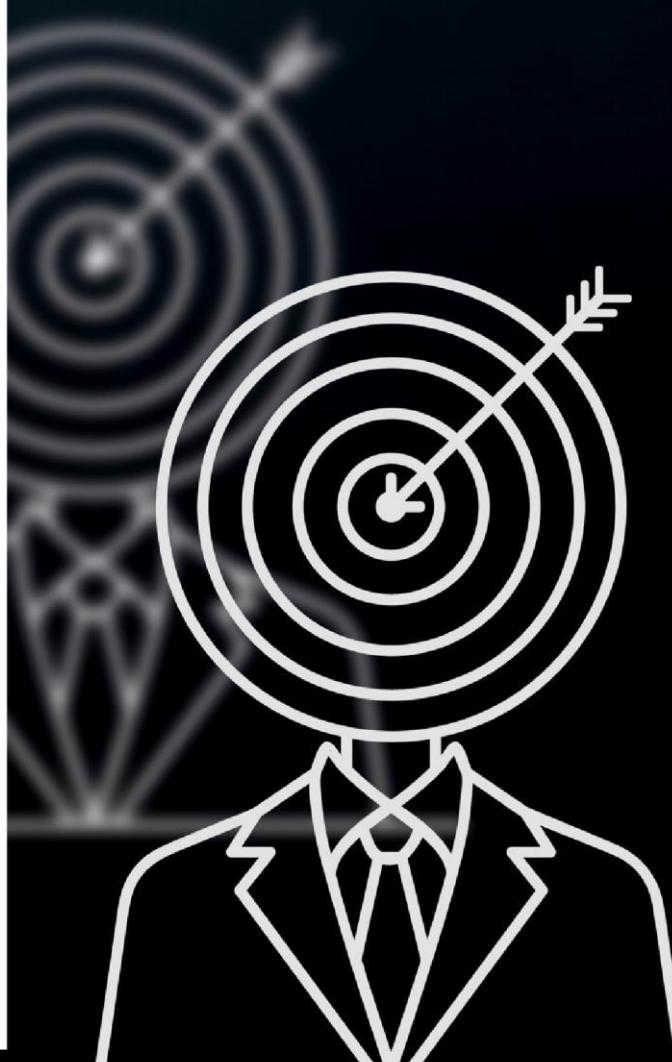


THE METHOD

CHAPTER 7: Researching your target audience: Why do people do what they do?

By Denise Goodwin, Kim Borg, Fernanda Mata,
Mark Boulet, Sarah Kneebone

www.behaviourworksaustralia.org



INTRODUCTION

Several years ago, we were assisting an Australian State Government Environment Department in the development of an electronic waste (e-waste) policy. They were interested in designing behaviour change interventions that would support more people in taking broken and unwanted electronic items to waste recovery centres for recycling. Yet they were unsure of what should inform the design of these interventions. Was more information about e-waste needed? Were social marketing campaigns required to persuade people of the benefits of recycling? Or were incentives needed to encourage more people to recycle?

To find out, we needed to identify the influences on the e-waste recycling behaviour. In particular, we were interested in the drivers and barriers to people taking e-waste to recovery centres. A Deep Dive was therefore required, but what research tools would be most suited for the task, and how might they best be used?

In the previous chapter, we suggested that you might not be “normal” - namely that what motivates, or prevents, your own engagement with a particular behaviour might be very different to what influences your target audience. We introduced a range of conscious and unconscious behavioural influences and argued that it is important to identify which are most relevant to the target audience for any particular behaviour, so that we can use those influences to design more effective behaviour change interventions.

Yet how do we find out more about the target audience, especially with regards to the relevant influences that might motivate, or prevent, them from engaging in a priority behaviour we want to promote? Choosing an appropriate research method to better understand influences is an important part of the Deep Dive component of the BWA Method. This chapter describes a number of methods that BWA researchers commonly use, identifies their benefits and limitations, and offers practical tips for successful implementation.

WHY DO WE NEED TO RESEARCH THE TARGET AUDIENCE?

A community, or the general public, is not a coherent or a homogenous group. It actually encompasses a huge amount of diversity and difference both within, and between, individuals of a group. These characteristics in turn can influence how, and why, a group participates, or not, in a priority behaviour.

The more we understand the target audience, the better we can explain what motivates people to engage in a behaviour, or what prevents them from doing so. For example, if we find out that the target audience has low levels of disposable household income, this might present a significant barrier to engaging in energy-saving behaviours that involve a financial investment, like installing solar panels or a solar hot water system.

At the same time, we can also see whether the priority behaviour is appropriate to the group we are trying to engage with. They may already be carrying out the priority behaviour, or it could be irrelevant to their lives. Is it worth promoting the behaviour of installing a rain-water tank if we find out that a large percentage of the target audience lives in apartment blocks?

Spending time researching the characteristics of the target audience, as well as their relevant influences in relation to the priority behaviour, means that we can design more appropriate, targeted and effective interventions.

What do we need to find out?

Considering our household electronic waste example, some of the information we might collect in a Deep Dive research process could include:

- Basic demographics: Age profile, gender, education levels, income, employment, occupation and cultural background.
- Household characteristics: Number of people living there, how many children, the type and size of the house.
- Life experience and psychosocial measures: Personal history, values, social norms, previous participation in the target behaviours, attitudes towards e-waste/recycling, and knowledge about the topic (e.g., recycling, collection site locations).

In addition to these internal (personal) and external (structural) characteristics, we can also investigate which behavioural influences might be most relevant to the priority behaviour from the perspective of the target audience. For example, whether attitudes towards the behaviour play a large role in deciding whether the target audience engages in it, or is participation based more on perceptions of social norms, or on things completely out of their control, like environmental contexts or opportunities provided by infrastructure.

HOW CAN WE RESEARCH THE TARGET AUDIENCE AND FIND OUT WHAT WE NEED TO KNOW?

This chapter describes some research methods we typically use in the Deep Dive phase of the BWA Method: questionnaires, interviews, observation, focus groups and mixed methods, including examples of where and how we have applied them. However, the first step for any research project, including uncovering audience insights, is to identify the research question(s):

Framing questions should always be the first step in the process. It should always be questions first, methods later. It makes no sense to decide “I am going to use questionnaires, interviews or observations” before clarifying the questions you wish to address (Wellington, 2000).

Identifying what you want to find out (i.e., what research questions you have), or what hypotheses you want to test, is fundamental to successful research design. It is vital to understand what you want to know before selecting the appropriate type of research method. For example, if you are interested in investigating the frequency of participation in a priority behaviour, quantitative data collection approaches, such as questionnaires, may be useful. To understand behavioural influences, such as norms and attitudes, qualitative data collection approaches, such as interviews, are more appropriate. If your project includes both types of questions, a mixedmethods approach, collecting the quantitative and qualitative data, should be used.

Selecting which methods to use (and when) should therefore be determined primarily by the problem you are addressing and the research questions you are interested in. Your research design will also depend on factors including what is already known, the size of the target audience, project budget, project deadlines, and other resource constraints. Unfortunately, there is no perfect research design solution to collect the perfect data. Each research method presents its own benefits and challenges which should be considered before making a decision. Tables 1 to 4 provide a quick snapshot of some of the pros and cons of key audience research methods, including when they can best be used. We then embark on some more in-depth accounts of individual methods, with accompanying case studies and practitioner tips.

Table 1: The pros and cons of questionnaires

Questionnaires: Face-to-face, online or hard copy	
Pros	Cons
Can get large numbers and representative samples fairly cheaply	Samples must be carefully selected to ensure they are statistically meaningful
Quick, particularly if running online, and easy to administer	Requires a separate data-entry step; coding of responses
Able to record a wide range of data – from demographics to attitudes, beliefs and participation rates	Dishonest or inaccurate answers
Standardisable	Subject to misinterpretation, depending on how questions are designed and asked

Source: Ruane (2005)

Use to: Collect more quantitative data and measure demographics, household characteristics, some psychosocial and life experiences. Investigate behavioural influences including attitudes, social norms, capability, opportunity, emotion and habits.

Table 2: The pros and cons of interviews

Interviews: Contact with participant in person, over the phone or email, with structured, semi-structured or open questioning	
Pros	Cons
Easier to ask open-ended questions, use probes and pick up on nonverbal cues	(Relatively) more expensive and slow data collection and analysis

Allows in-depth exploration of responses and pursue “hunches”	Requires strong interviewing skills
Line of discussion can be tailored to the individual	Reactive effect; interviewer’s presence and characteristics may bias results
Easier to reach specific individuals with a personalised approach	Respondents who prefer anonymity may be inhibited by personal approach

Source: Yeo, Legard, Keegan, Ward, McNaughton Nicholls, & Lewis (2014).

Use to: Collect more qualitative data around demographics, household characteristics, some psychosocial and life experiences. Investigate behavioural influences including attitudes, social norms, capability, opportunity, emotion and habits.

Table 3: The pros and cons of observations

Observation: Observe how beliefs and practice reality and explore beliefs, behaviours, and a		inconsistencies in reported
Pros	Cons	
Setting is natural, flexible and unstructured	Requires skilled observer	
Evaluator may actively participate or observe passively	Observations cannot be generalised to entire population unless a plan for representativeness is developed	
Can be combined with a variety of other data collection methods	Difficult to determine root cause of observed behaviour	
Most useful for studying a small unit such as a classroom, council, etc.	Difficult to operationalise and code behaviours	
Reduce bias; e.g., recall, social desirability	Hawthorne effect - if group is aware they are being observed, resulting behaviour may be affected	

Source: McNaughton Nicholls, Mills, & Kotecha (2014).

Use to: Collect qualitative or quantitative data to understand the context and environment within which the priority behaviour is occurring, the frequency of behaviour engagement (sometimes), and to identify behavioural compliers and non-compliers. Can also help to identify influences such as social norms, habits, emotion, biases, capability and opportunity.

Table 4: The pros and cons of focus groups

Focus groups: Invited panels of stakeholders or target a facilitator	audience members discuss specific questions posed by
Pros	Cons
Good way to gain multiple perspectives around an issue	Social effects may mean that some voices are drowned out; inequitable representation of participants and false agreement on issues
Discussions can explore more nuanced questions and generate rich data and understanding	Produces large amounts of qualitative data that can be challenging to analyse
Works well with specifically selected or invited groups	Can be expensive and time consuming to set up and run - particularly if there are key individuals that need to be included.

Source: Center for Disease Control and Prevention (2008)

Use to: Collect qualitative data around behavioural influences, including attitudes, social norms, capability, opportunity, emotion and habits.

Survey research

In the social and behavioural sciences, one of the most common research methods is asking lots of people questions about themselves and their experiences to understand the target audience. This is known as **survey research**. The way these questions are delivered to audiences can take different forms, but the main methods are **questionnaires** and **interviews**.

Questionnaires

CASE STUDY

Research Problem

Australia Post wanted to conduct some Deep Dive audience research to better understand digital inclusion in Australia (Borg & Smith, 2017). The target audience was Australian adults who did *and* *did not* use the internet. We wanted to hear from a large cross-section of the Australian population, so we needed a research method that could reach a large number of people. But we couldn't reach our audience or administer our questions online.

Research Design

We used random digit dialling (RDD) (i.e., randomly selecting and dialling phone numbers) to reach the target audience. The questionnaire was then administered via computer assisted telephone interviewing (CATI) (i.e., interviewers read the questions out and record people's answers). This approach allowed us to reach a large (over 1,500 people) and representative sample of Australians (i.e., from each state/geographic location, from a mixture of age groups and genders).

Questionnaires are an instrument used to gather specific information from a large number of people (Ruane, 2005). The questions can be “open” (i.e., describing something in their own words) or ‘closed’ (i.e., selecting from a list of options or within a range). Closed questions can capture scale numbers (e.g., “what is your height/weight?”), ordinal responses (e.g., “from 1=Strongly disagree to 5=Strongly agree, how much you agree or disagree with X?”), or categorical responses (e.g., “what is your favourite colour?”).

In addition to designing the instrument, it is important to consider how to deliver the questions to your target audience - i.e., what is your survey approach? For example, who is your target audience (e.g., farmers)? How can you reach your target audience (e.g. do you have access to a list with everyone’s contact details)? How will you administer the questions (e.g., online, in person, phone, mailout)? How many people do you need to answer your questions (e.g., 20 vs 5,000)? The answers to these questions will determine how you conduct your survey.

In the digital inclusion example, computer assisted telephone interviewing helped us reach a large number of people who used and did not use the internet. However, we also didn’t want to exclude vulnerable populations, such as those experiencing homelessness, who may not have a phone. So we also administered some questionnaires via in-person interviews. Surveying via interviews can be useful, but it is also quite time and resource intensive to administer. The issue of “social desirability” is also higher when an interviewer is involved (i.e., the tendency to respond in a manner perceived as favourable to others). For these (and other) reasons, most surveys are conducted using self-administered questionnaires via an online platform or hard-copy paper and pencil.

There is still the question of how to reach the target audience with self-administered questionnaires. Some methods include “convenience sampling”, where participants are recruited via existing networks (such as lists of members, or social media). “Snowballing” techniques can also be used to increase recruitment reach (i.e., asking others to also share the instrument with their networks). The downside of convenience sampling is that your final sample might not “look like” or be representative of your target audience.

For example, the Victorian Department of Environment, Land, Water and Planning wanted to know how

Victorians value and act to protect nature. If they had surveyed their own networks, it is likely that the responding sample would have been pretty “green” already. Instead, the Department engaged BWA to conduct the research for them with a representative sample of over 3,000 Victorians (Meis-Harris et al., 2019). To reach such a large sample, we worked with a research company that curates their own panel of participants (by constantly recruiting people from all walks of life). This enabled us to survey people from all over the state, including those who are very connected to nature and those who are entirely disconnected.

By surveying large and diverse samples, we learned from both the “Digital Inclusion” survey and the “Victorians Valuing Nature” survey that, not only were our audiences not like us, but they weren’t like each other either. Australians engage with the internet in different ways and for different reasons; some people preferred socialising and playing online games, while others preferred practical activities like searching for information or online banking (see

Chapter 3). Similarly, Victorians were connected to nature in different ways, spent time in nature differently, and engaged in different protective behaviours, with women and older people reporting greater connection to nature than men and younger people.

Practitioner tip!

When designing a survey, start from the end – what do you want to be able to say? What research questions do you want to answer? How will you analyse your data? Do you or someone in your team have the capacity to conduct complex statistical analyses? Or are you more interested in the average score or percentage of people who select a particular response? Once you know the types of answers you’re looking for, you can start to operationalise your questions, and determine if a survey is the right tool for the job. Piloting your survey (with 5 - 10% of respondents) helps to check for ambiguity and confusion over questions before running it with your full respondent list. **Interviews**

CASE STUDY

Research Problem

In partnership with a clinical team at the Victorian Spinal Cord Service, we wanted to narrow the gap between current practice and those advised in clinical practice guidelines (Nunn et al., 2018). Specifically, we wanted to explore the drivers and barriers associated with transferring patients following an acute spinal cord injury (SCI) from an indwelling catheter (IDC) to an intermittent catheter (IC), as this has been shown to reduce the incidence of urinary tract infections. These are a common and costly complication of SCI.

Research Design

We used interviews, rather than questionnaires, to gather this information because it is a relatively small group and there was limited knowledge on the drivers of behaviour of the target audience. Interviews also enabled the research team to understand contextual factors that influenced behaviour in this particular hospital setting. The aim of the project was to design and implement a behaviour change strategy to optimise the early use of ICs in newly acquired SCI individuals.

If you have the time and resources, one example of a more in-depth approach to understanding behavioural drivers and contextual information to support intervention implementation is to conduct qualitative **interviews**. In essence, interviews are a method of data collection that involves two or more people exchanging information through a series of questions and answers. Interviews are usually semi-structured, meaning questions are openended and offer participants an opportunity to provide a depth of information. Semistructured interviews are used extensively as an interviewing format with individuals or even with a group (referred to as focus groups).

Interviews can help you explain, better understand, and explore people's opinions, behaviour, experiences and phenomenon. They also provide the interviewer an opportunity to explore emerging themes as well as salient issues in relation to the development and understanding of a problem. They enable us to dig a little deeper into the story behind responses you may have from a survey.

Interviews can be an important part of the Deep Dive phase of the BWA Method, to inform the development and implementation of behavioural strategies. We often use interviews to understand drivers of behaviour by asking broad questions relating to the barriers and enablers of a particular behaviour. In our spinal example, we explored clinicians' and patients' views on SCI urinary catheter care. Thematic analysis was guided by the Theoretical Domains Framework (TDF - see Chapter 6 for a summary) to categorise identified behavioural drivers of SCI urinary catheter care. Findings from the TDF analysis showed it wasn't just a knowledge and skills gap driving behaviour, although these areas did need to be addressed. There were also "social influences", "environmental context and resources", "beliefs about capabilities" and "beliefs about consequences" that were found to drive behaviour.

Knowing these drivers of behaviour enabled the research team to identify behaviour change strategies that were more likely to support clinical practice change. The project team engaged with nurses and clinicians in the hospital, a local peer support network, and people living with SCI, to co-design interventions to enable implementation and support knowledge translation. Figure 1 illustrates the dominant behavioural drivers mapped to the TDF and the behaviour change strategies. After a three month audit, the number of days to IDC removal and the start of ICs reduced from a median of 58 days to 19.5 days. Readmittance to the acute ward from the rehabilitation unit (6 months following an acute SCI) also significantly reduced, with only one case recorded since the intervention was implemented.

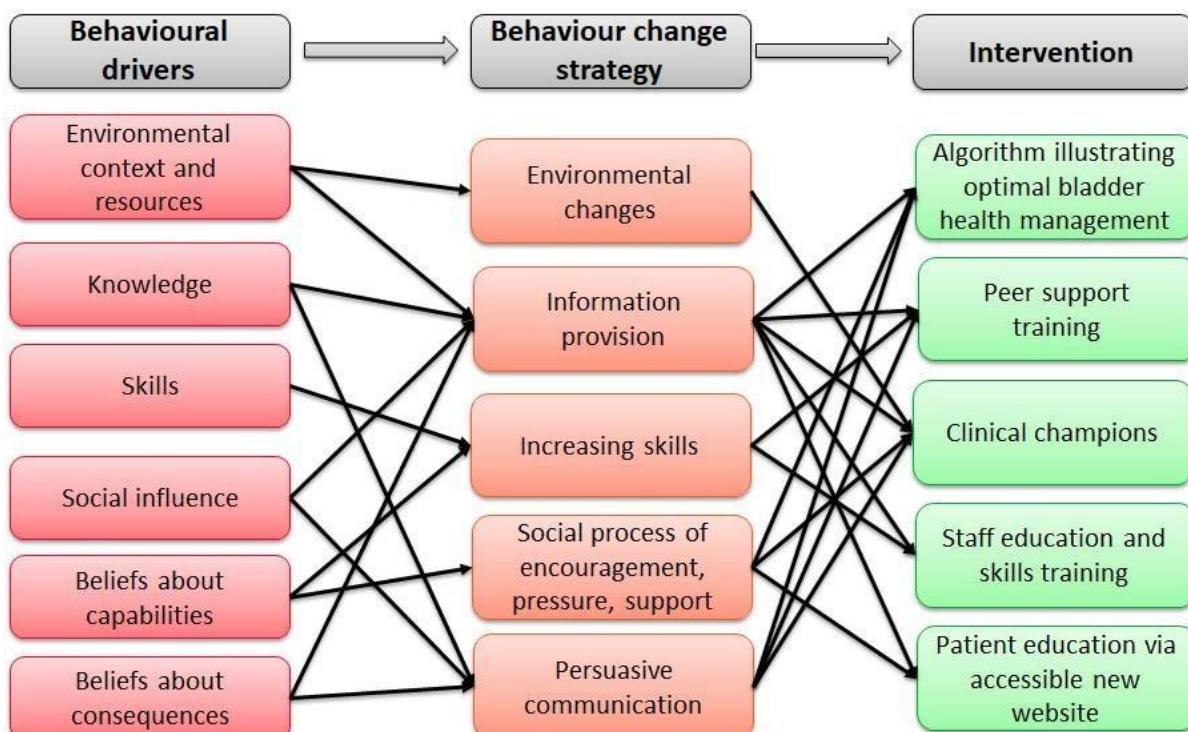


Figure 1: Dominant drivers of behaviour identified by the TDF, and associated behaviour change strategies to guide a multifaceted intervention to support early transition to IC following acute SCI.

Practitioner tip!

Conducting successful interviews is an art which takes time to develop. To get started, keep your interview guide short - around 5-7 questions is usually enough. It is important to avoid

questions that can be answered with a simple yes or no. Instead ask short, jargon-free questions that will elicit a detailed response. Listen to what your respondent is saying - don't get caught up in thinking about your next question or what you want to cover. Remember this is their story and usually when things go "off track", this is where the real detail can be found! Use active listening skills to show you are engaged. This can be trickier over the phone, but using affirmations like "Oh"; "Hmm"; or "I see" provides the participant cues that you are engaged. And finally, practice, practice, practice! Remember to try your questions out first to check the length of interview, clarity of instructions and question structure.

Observation

CASE STUDY

Research Problem

A food rescue organisation was interested in understanding more about what motivates food use behaviours in the home, and how food waste might emerge from these behaviours. We wanted to gain better insights into how people perform, or experience, different food related behaviours (like shopping and cooking).

Research Design

As we were particularly keen to see how these behaviours "played out" in their actual household contexts, rather than a more abstracted recollection of these, we could not rely on traditional survey or interview responses. If we were going to design interventions that would change behaviours in people's kitchens, we needed to better understand this context. We therefore used an observational style research method to uncover these particular insights.

The majority of existing food waste research tends to rely on the use of survey-based questionnaires. While these are an effective way of gathering information from a large sample of people about what they do and think with regards to an issue or behaviour, they can suffer from limitations associated with self-report bias (see Table 1). While they capture *what people say they do*, they do not provide much insight into *what they actually do*. They also can miss the unconscious influence of habits, biases and the physical environment on behaviours.

Observational research methods are essentially about collecting data by watching. Traditionally, this has involved the researcher being physically present to observe participants engaging in particular activities in particular contexts. This gives a truly front-row seat to the action and fully embeds the observer in the physical and social realities of the target audience.

While this approach can be a little resource intensive, and the presence of the researcher can artificially influence the behaviours of participants, technology has enabled new methods that are less intrusive and more efficient. We show two examples in this section.

For our food waste study, we drew on the use of *participant generated videos* to conduct observational research. Using a marketing company that specialises in video-based research, a panel of participants was recruited and set a series of cooking and other kitchen tasks. Participants had already been trained by the marketing company to use their smartphones to record videos and recorded themselves performing the different tasks. These

were then uploaded to an online platform and participants recorded themselves again, this time watching their initial videos and describing the different things that influenced particular actions or decisions that were evident on screen (see Figure 2).



Figure 2: Example of the dual-frame videos created in this study. The main image is from the first video the participant records of their performance of a provisioning practice (unpacking from a food shop). The smaller image is from the second video the participant records while watching the first video and reflecting on their actions. Face blocked for participant privacy.

This provided incredibly rich data! Not only did we have access to the initial video that allowed us to observe the actions of the participants, but we also had the explanations of the participants themselves. The visual medium, and participant control of the camera, gave a strong sense that we were “looking at” what participants found important when explaining particular actions and choices, rather than imposing our own assumptions of what motivated them. It allowed us to get one step closer to consumers’ food related behaviours in context, and to observe the complexities, trade-offs and messiness that accompany these behaviours and that lead to food waste. We also were able to gain insights into how the physical environment of the kitchen influenced behaviours. All this would have been missed if we had relied on traditional survey or interview formats which can lead to more “sanitized” accounts from respondents.

A second observational method we have utilised at BWA is the use of eye tracking technology. Between 2013 –

2016, the Victorian Environment Protection Authority partnered with BWA, Lifesaving Victoria and Federation University to investigate new communication channels to ensure that beach users are aware of the water quality prior to visiting the beach. One study involved BWA redesigning its signage around Port Phillip Bay beaches to better communicate water quality in real time. Using Monash University’s Behavioural Research Laboratory, a series of signs with varying content were tested for their impact on attention paying to content and assimilation of the information contained. An eye-tracking computer captured and measured where study participants looked (and for how long) within the images shown (see Figure 3). The observations found that signs that used symbols (such as smiley faces) were noticed more often than signs that relied on just colour and arrow indicators.



*Figure 3: Example of the two signs *in situ* and the computer-based eye tracking, which identifies areas of a visual field brought into high resolution by the observer's focus.*

Practitioner tip!

Observational research can be a little overwhelming at first. You are not just recording what people are saying, but also what they are doing. This generates a very large and diverse data set, which can be analysed in a number of ways. While you don't want to close yourself off to new, emergent and unexpected findings, having some sense of what you might be looking for (e.g., by using some sort of analytical framework) beforehand can help to navigate the large amount of data and prevent you from drowning. Remember to pilot your method by trialling your observation approach *in situ* to check the clarity of your behaviours of interest and address any issues around coding responses in data collection.

Mixed-methods

CASE STUDY

Research Problem

The Victorian Department of Families, Fairness and Housing (DFFH, formerly the Department of Health and Human Services) wanted to understand how to mobilise the wider community to offer support to vulnerable families in Victoria (i.e., families who experience issues that could put children at risk of harm, like poverty, homelessness, mental health issues, domestic/family violence, etc.).

Research Design

This multifaceted problem required a multi-stage mixed methods research project. After initially running a literature and practice review (see Chapter 1), we conducted focus groups with people in Melbourne and regional Victoria. The aim of the focus groups was to understand the *types* of drivers and barriers for engaging in supportive behaviours, including the two priority behaviours (see Chapter 5): listening and talking to vulnerable families, and participating in community activities. Insights from the focus groups then helped us design the questionnaire for a state-wide online survey of over 5,000 Victorians. The aim of the survey was to understand what was happening at the state level (e.g., were particular

communities more or less supportive?) and to determine *the extent to which* the drivers and barriers identified in the focus groups influenced the supportive behaviours.

A **mixed-methods approach** involves collecting, analysing, and “mixing” qualitative (e.g., interviews, focus groups, observation) and quantitative research (e.g., surveys) into a single study. This research approach aims to provide a framework for integrating quantitative and qualitative methods, and is commonly used for gaining a better understanding of the target audience. When used in combination, qualitative and quantitative methods complement each other, and enable researchers to better understand their target audience through different lenses. The rationale for using a mixed-methods approach is based on the understanding that neither qualitative nor quantitative approaches are fully able to address all the details and nuances of the target audience.

Sometimes, the findings of a study can provide an incomplete understanding of the target audience and further explanation is required. A common scenario is when quantitative findings require further explanation as to what they mean. It’s also possible that the researchers want to examine a subset of the target audience in greater depth after conducting a survey with a large sample of the population. In these cases, a mixed-methods approach is recommended, with the second study approach improving the understanding of the first. Quantitative findings can help us capture the frequency of behaviours and establish the relationships between variables, but the more detailed understanding of what each of these relationships mean is lacking. Qualitative data can enhance that understanding by bringing together context and, therefore, more information about the behaviours and attitudes of the target audience. As such, mixed-method approaches can promote different avenues of exploration that enable researchers to better understand their audience.

The mixed-methods approach is also useful when you know very little about the target audience and you need to better understand who they are before starting the quantitative phase of a study. In this case, it is common to conduct focus groups or interviews to gain an initial understanding of the target audience, and then use the findings of this phase of the research to inform the development of a survey to be distributed to a larger group of participants. This quantitative phase of the study can help researchers understand frequencies of behaviours of interest, and how attitudes and perceptions identified in the qualitative phase of the study can influence specific behaviours.

Practitioner tip!

The decision to use a mixed-method approach to better understand the target audience is based on the value that using both methods (qualitative and quantitative) has compared with using a single method. However, it’s important to keep in mind that the mixed-methods approach can significantly increase the complexity of conducting your research project. This is because collecting and analysing both qualitative and quantitative data within the same study often requires more resources, including time and personnel, as well as high-level understanding of different research methods (qualitative and quantitative) and data analysis approaches.

CONCLUSION

Back to our introductory e-waste case study. To explore what potentially influences the intention of someone to take e-waste to a waste recovery centre for recycling, we ended up taking a mixed-methods approach:

- **We interviewed** a small sample of people (about 10) to identify some of their beliefs about the advantages and disadvantages of the behaviour (their attitudes), who would approve or disapprove of the behaviour (perceived social norms) and what might hinder or help them to perform the behaviour (perceived behavioural control). These interviews allowed us to engage in-depth with participants about their different beliefs, to probe them for more information and to capture unexpected responses.
- **We then used an online questionnaire** with a larger sample (about 300) to measure the strength of the beliefs identified in the interviews. Namely, of the different attitudinal, social norm and perceived behavioural control beliefs mentioned in the interviews, which ones were more likely to predict someone's intention to recycle e-waste? The survey allowed us to engage more broadly across the population and gather a larger sample for more robust, representative, findings.

These two methods combined helped to provide direction to the Environment Department's e-waste policy and what should be prioritised. We found low knowledge of what "e-waste" actually meant and recommended an information campaign that explains the term to people. We also found that people were more likely to recycle ewaste if they knew what could be recycled (capability) and where they could go (opportunity). We therefore recommended ensuring a consistent, and easily accessible spread of recycling sites across the region and a consistency between them in terms of the items that were accepted.

To summarise, this chapter has shared some of the different research methods that BWA researchers often use to understand more about a target audience and to explore different motivators, or barriers, to target behaviours from their perspectives. As always, we are only able to give a brief overview of some of the more common methods and recommend engaging with the additional resources provided with this chapter to learn more about the range of possible research methods and their application.

We leave you with one final thought when it comes to selecting, and using, research methods. It is easy to become overwhelmed by what is available to use and to become overly concerned with the pros and cons of each method. While these issues are endlessly fascinating to university-based academics, and countless books and articles have been written about them, at the end of the day, *anything* that gets you one step closer to understanding a behaviour from the perspective of your target audience is better than going with your own assumptions. Even a short informal chat about your target behaviour with one or two people from your target audience is going to lead to a more appropriate and tailored intervention than if you were to stay at your desk and develop policy or programs based on what you *think* will influence people.

And that is what our next chapter will focus on - using the insights gained from the application of different research methods to design interventions that change behaviour.

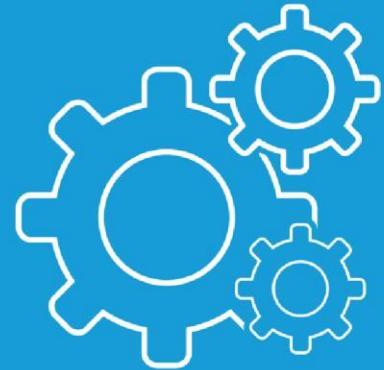
REFERENCES AND OTHER SUGGESTED READING

- Bernard, H. R., & Russell, H (2013). *Social Research Methods: Qualitative and Quantitative Approaches*. SAGE.
- Borg, K., & Smith, L. (2017). Digital Inclusion: Report of Online Behaviours in Australia 2016.
https://auspost.com.au/content/dam/auspost_corp/media/documents/research-reportdigitalinclusion.pdf
- Boulet, M., Hoek, A., & Raven, R. (2021). The gaze of the gatekeeper: Unpacking the multilevel influences and interactions of household food waste through a video elicitation study. *Resources, Conservation and Recycling*, 171, 105625.
- Center for Disease Control and Prevention. (2008). Data collection methods for program evaluation: Focus groups. Retrieved October, 27, 2008.
- Creswell, J. W. (2013). *Research Design: Qualitative, Quantitative, And Mixed Methods Approaches*. SAGE.
- Creswell, J. W., & Clark, V. L. P. (2017). Chapter 3, Choosing a mixed methods design. In *Designing and Conducting Mixed Methods Research*. SAGE.
https://www.sagepub.com/sites/default/files/upm-binaries/35066_Chapter3.pdf
- Donaldson, S. I. & Grant-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. *Journal of Business and Psychology*, 17(2), 245-260.
- McNaughton Nicholls, C., Mills, L. & Kotecha, M. (2014). Observation. In Ritchie, J., Lewis, J., Nicholls, C. M., & Ormston, R. (Eds.), *Qualitative Research Practice: A Guide for Social Science Students and Researchers*, 243-268. SAGE.
- Marsden, P. V. & Wright, J. D. (2010). *Handbook of Survey Research*. Emerald Group Publishing.
- Meis-Harris, J., Saeri, A., Boulet, M., Borg, K., Faulkner, N., & Jorgensen, B. (2019). Victorians Value Nature: Survey Results.
https://www.ari.vic.gov.au/_data/assets/pdf_file/0030/443379/Victorians-ValueNaturesurvey-results-report-2019.pdf
- Nunn, A., Bragge, P., Goodwin, D. & Byrne, C. (2018). Output Report: Optimal bladder health following Spinal Cord Injury. The Institute for Safety Compensation and Recovery Research.
- Ruane, J.M. (2005). Chapter 9, Our inquisitive nature: The questionnaire. In *Essentials of Research Methods: A Guide to Social Science Research*, 123-145. Blackwell Publishing.
- Rickinson, M., (2012). *ELT Research Guides: Effective Research Interviews*. Oxford University Press
- Rickinson, M., (2013). *ELT Research Guides: Effective Observational Research*. Oxford University Press

Ritchie, J., Lewis, J., Nicholls, C. M., & Ormston, R. (Eds.). (2014). *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. SAGE.

Wellington, J. (2000). *Educational Research: Contemporary Issues and Practical Approaches*. Bloomsbury Publishing.

Yeo, A., Legard, R., Keegan, J., Ward, K., McNaughton Nicholls, C., & Lewis, J. (2014). Indepth interviews. In Ritchie, J., Lewis, J., Nicholls, C. M., & Ormston, R. (Eds.), *Qualitative Research Practice: A Guide for Social Science Students and Researchers*, 177-210. SAGE.



THE METHOD

CHAPTER 8: From insights to interventions

By Lena Jungbluth, Kun Zhao, Breanna Wright,
Bernice Plant and Denise Goodwin

www.behaviourworksaustralia.org



INTRODUCTION

A large number of hospital deaths in Australia are preventable. These deaths are often preceded by abnormalities in vital signs and other observations. Early identification of deterioration can assist in providing earlier and lower-level intervention to patients. However, despite national standards for the recognition and response to deterioration, there is evidence that warning signs of patient deterioration are not always recognised or followed up.

To address patient safety in Victorian public hospitals, the Victorian Managed Insurance Authority (VMIA) saw an opportunity to improve how doctors and nurses were having conversations about escalation of care—when a patient requires additional treatment. This is a multi-faceted problem that involves a lot of different audiences and stakeholders. There is the patient and their family or caregivers, nurses and midwives, and doctors (from junior doctors to registrars and consultants). All of these audiences have a role to play in escalation of care conversations.

Partnering with VMIA and stakeholder groups, BehaviourWorks Australia (BWA) embarked on a project to come up with behavioural strategies to improve escalation of care for patients experiencing deterioration in the hospital setting. Reaching the end of our Deep Dive phase of the BWA Method, we had gathered a number of insights on the behavioural influences impacting medical professionals, as well as findings from our evidence and practice review. Our next step was to bring these together and develop a shared understanding among stakeholders in order to design an intervention that could be trialled and scaled across Victorian hospitals.

When trying to change people's behaviour, it is common to jump to "default" solutions, like education or offering incentives. This is because people often make assumptions about the types of interventions that will be effective. For example, a commonly-held belief is that when someone is aware of what they should do, they will simply do it—we just need to provide them with the necessary information. But if, like us, you can reflect on previous attempts at achieving your New Year's resolutions, you'll recognise that having the necessary knowledge is not sufficient for us to stick to our resolutions. Even when we know that eating too much chocolate cake or drinking alcohol is not healthy for us, we may still do it. This is because, in the case of us eating too much chocolate cake (hypothetically...), a lack of knowledge is not the key barrier to us stopping, and thus no amount of education will help us stick to our resolutions in this case. This is a reminder of the importance of developing intervention options that align with any insights you have about the problem.

In the previous chapter, we introduced a range of methods that can be used during the Deep Dive phase to understand a target audience, such as questionnaires, interviews, observation and focus groups. Such insights, combined with those often gathered from evidence and practice reviews (see Chapters 1 and 4) can inform interventions to change the behaviour of a target audience. While this may sound like a relatively simple task (i.e., insights = intervention = behaviour change), there is a process to step through and a number of factors that should be considered.

In this chapter, we take you through the final leg of the Deep Dive phase of the BWA Method, where you will go from understanding the drivers and barriers of your target behaviour to developing interventions that reflect these insights. Here, we define a *behaviour change intervention* as any activity designed to change a particular behaviour(s)

of a specific audience(s) to help address a social problem, and which draws on behavioural science principles (see also the definition used in Michie, van Stralen, & West, 2011). This is not to be confused with the method or mode by which an intervention is delivered (e.g., a mobile phone app, a workshop, or a poster)—we would consider these the intervention *mode of delivery* (Michie, Atkins, & West, 2014).

This chapter is intended to provide you with some key steps and considerations when going from insights to interventions. While it is not a comprehensive guide on intervention design, our [online resources](#) offer some additional recommendations to consider. Furthermore, behavioural problems are often complex (in particular, see wicked problems in Figure 1, Chapter 5) and influenced by multiple players in a system (as we learned through systems mapping in Chapter 2). As a result, there are two important caveats to bear in mind when designing your behaviour change intervention.

First, taking the Deep Dive approach to developing intervention options does not, of course, guarantee that your intervention will be successful. However, designing and selecting an intervention that addresses key barriers (or leverages off key drivers) of the behaviour will increase your chance of success. Second, targeting a single behaviour limited to one audience may not be sufficient, and a combination of interventions addressing multiple individual behaviours operating at multiple levels of the system may be needed. With this in mind, this chapter focuses on the basic building block of behaviour change: designing a suitable intervention for a single target behaviour and audience.

You'll notice throughout this chapter that we draw a lot from the work of Professor Susan Michie at the Centre for Behaviour Change at University College London and her colleagues. The frameworks they developed have influenced our practices (as well as many other people working in behaviour change!) and dominate the approaches we use in most projects. Where relevant, we also point out some other frameworks that you might find helpful.

WHAT TYPES OF INTERVENTIONS EXIST?

Chapter 6 introduced us to a range of different influences on our behaviour and Chapter 7 covered a number of methods for understanding these influences, which can then guide our intervention design. Otherwise, by simply following our assumptions, as mentioned in the introduction of this chapter, we risk designing (and paying for) yet another education or awareness intervention which doesn't achieve the desired behaviours.⁵ For example, hand hygiene among healthcare workers is a core behaviour used in hospitals to reduce the incidence of healthcare-associated infections that can have both health and economic consequences. Yet, despite repeated education and persuasive campaigns over the past two decades promoting hand hygiene in healthcare settings (including the World Health Organization - WHO, 2009), a recent Cochrane Review revealed that this "simple" behaviour was not fully followed by healthcare workers globally (Gould et al., 2017). Clearly, we need

⁵ This of course doesn't mean we should exclude education from the list of possible intervention types for our target audience. Education can play a very important role for the uptake of behaviours in instances where our audience actually lacks knowledge or awareness about a behaviour or about certain aspects of performing a behaviour.

to look to other types of interventions to improve hand hygiene practices among healthcare workers.

Beyond education, what other types of interventions could we draw on? One way to begin your intervention design is to lay out in front of you all the different intervention types that exist. A research team led by Professor Susan Michie has identified nine *intervention functions*, based on a systematic review consisting of 19 different behaviour change frameworks (Michie et al., 2011; 2014)⁶. This classification is based on the “function” of the

intervention, or, as the authors describe, the “broad categories of means by which an intervention can change behaviour” (Michie et al., 2014, p. 109).

Some of these intervention functions are very well known and have long dominated the practice of behaviour change, such as education (e.g., awareness campaigns in public health) and coercion (e.g., speeding fines). Other intervention functions have begun to receive greater interest more recently, such as environmental restructuring. Many nudgetype behavioural interventions, including the use of defaults or salient displays of messages, are examples of this kind.

An overview of these nine intervention functions is provided in Table 1, along with some examples of each function. Together, these intervention functions serve as a “toolkit” when thinking about the different types of interventions that could be developed for a given behavioural problem.

Table 1: Overview of intervention functions (Michie et al., 2011; 2014)

INTERVENTION FUNCTION	DESCRIPTION	EXAMPLES
Education	Increasing knowledge or understanding	Providing information to promote healthy eating
Persuasion	Using communication to induce positive or negative feelings or stimulate action	Using imagery to motivate increases in physical activity
Incentivisation	Creating expectation of reward	Using prize draws to induce attempts to stop smoking
Coercion	Creating expectation of punishment or cost	Raising the financial cost to reduce excessive alcohol consumption
Training	Imparting skills	Advanced driver training to increase safe driving
Restriction	Using rules to reduce the opportunity to engage in the target behaviour (or to increase the target behaviour by reducing	Prohibiting sales of solvents to people under 18 to reduce use for intoxication

⁶ Another behavioural taxonomy includes MINDSPACE (Dolan, Hallsworth, Halpern, King, Metcalfe, & Vlaev, 2012), consisting of nine approaches in which behavioural science could be applied to public policy. The MINDSPACE framework is also an example of a general model in which behaviour can be influenced without the need for a targeted Deep Dive. Similar models are covered later on in this book.

	the opportunity to engage in competing behaviours) ⁷	
Environmental restructuring	Changing the physical or social context	Providing on-screen prompts for GPs to ask about smoking behaviour
Modelling	Providing an example for people to aspire to or imitate	Using TV drama scenes involving safe-sex practices to increase condom use
Enablement	Increasing means / reducing barriers to increase capability (beyond education and training) or opportunity (beyond environmental restructuring)	Behavioural support for smoking cessation, medication for cognitive deficits, surgery to reduce obesity, prostheses to promote physical activity

MATCHING TYPES OF INTERVENTIONS TO BEHAVIOURAL INFLUENCES

Now that we have a “toolkit” of different intervention functions in front of us, our next step is to identify which of these is best placed to address the relevant influences (i.e., drivers and barriers) of our target audience. But how do we know which intervention function is most appropriate for each behavioural influence?

Luckily for us, some of this matching work has already been done. You might remember that in Chapter 6 we introduced you to two frameworks for understanding and categorising behavioural influences during the Deep Dive. The Theoretical Domains Framework (TDF) identifies 14 general domains (or influences) on behaviour (Cane, O’Connor, & Michie, 2012), such as knowledge, skills and intentions. Alternatively, you may prefer to use the BWA Influences on Behaviour, which our team have adapted from Darnton (2008) and which consists of eight categories of influences such as attitudes and opportunity.

In both cases, the nine intervention functions have been mapped to behavioural influences in the two frameworks. For the TDF, Michie and her team (2014) mapped out which intervention functions were most promising in addressing each of the TDF domains, based on the consensus of a large group of behavioural experts. For the BWA Influences on Behaviour, our own team mapped intervention functions according to the behavioural influences they were most likely to change. This was a more applied mapping exercise that was based on our experiences and suited to the projects that we worked on.

In both cases, the mapping draws on comprehensive classifications of behavioural influences and helps pinpoint which intervention functions are worth exploring. These maps are provided in Tables 2 and 3, each of which is followed by a case study where they were used. Note, however, that the mapping—while developed by behavioural experts—refers to general links between behavioural influences and intervention functions. As this is not

⁷ As outlined in Chapter 5, we recommend identifying behaviours by framing them as positive or solution-oriented (e.g., “take uneaten food home” rather than “don’t put food waste in bins”). Similarly here, we’d recommend focusing on increasing the target behaviour by using rules to reduce the opportunity to engage in competing behaviours.

exhaustive (and there are always exceptions), it is important to complement your mapping of influences and intervention functions with other evidence, such as a review of the literature and stakeholder consultation (more on this later).

Table 2: Links between TDF and intervention functions, adapted from Michie et al. (2014)

DOMAIN	DESCRIPTION	INTERVENTION FUNCTION
Knowledge	An awareness of the existence of something	Education
Skills	An ability or proficiency acquired through practice	Training
Social / professional role or identity	A coherent set of behaviours and displayed personal qualities of an individual in a social or work setting	Education Persuasion Modelling
Beliefs about capabilities	Acceptance of the truth, reality, or validity about an ability, talent, or facility that a person can put to constructive use	Education Persuasion Modelling Enablement
Optimism	The confidence that things will happen for the best or that desired goals will be attained	Education Persuasion Modelling Enablement
Beliefs about consequences	Acceptance of the truth, reality, or validity about outcomes of a behaviour in a given situation	Education Persuasion Modelling
Reinforcement	Increasing the probability of a response by arranging a dependent relationship, or contingency, between the response and a given stimulus	Training Incentivisation Coercion Environmental restructuring
Intentions	A conscious decision to perform a behaviour or a resolve to act in a certain way	Education Persuasion Incentivisation Coercion Modelling
Goals	Mental representations of outcomes or end states that an individual wants to achieve	Education Persuasion Incentivisation Coercion Modelling Enablement
Memory, attention and decision processes	The ability to retain information, focus selectively on aspects of the environment and choose between two or more alternatives	Training Environmental restructuring Enablement

Environmental context and resources	Any circumstance of a person's situation or environment that discourages or encourages the development of skills and abilities, independence, social competence, and adaptive behaviour	Training Restriction Environmental restructuring Enablement
Social influences	Those interpersonal processes that can cause individuals to change their thoughts, feelings, or behaviours	Restriction Environmental restructuring Modelling Enablement
Emotion	A complex reaction pattern, involving experiential, behavioural, and physiological elements, by which the individual attempts to deal with a personally significant matter or event	Persuasion Incentivisation Coercion Modelling Enablement
Behavioural regulation	Anything aimed at managing or changing objectively observed or measured actions	Education Training Modelling Enablement

Back to our escalation of care case study. Our Deep Dive research revealed that there was a clear gap in the knowledge and skills of medical staff. Specifically, we found that nurses, midwives and doctors did not have the opportunity to develop skills in how to have conversations about escalation of care in the early stages of patient deterioration, particularly when dealing with barriers such as hierarchical culture and pushback. Looking at Table 2 which maps the theoretical TDF domains with intervention functions, we can see that appropriate interventions for addressing “Knowledge” and “Skills” are through “Education” and “Training”, respectively. This then helped guide our thinking about what sort of interventions within the category of education and training might be suitable solutions.

Table 3: Links between BWA Influences on Behaviour adapted from Darnton (2008) and intervention functions

INFLUENCE	DESCRIPTION	INTERVENTION FUNCTION
Attitudes	Overall favourable or unfavourable evaluation of engaging in a behaviour	Persuasion Education Incentivisation Coercion
Social norms	Social rules that indicate what the common, expected and acceptable behaviours are in a particular situation	Persuasion Modelling
Capability	Personal physical, financial or psychological capabilities to undertake the behaviour	Training Education

Opportunity	Factors or circumstances beyond the individual that provide them with the means to carry out the behaviour	Environmental restructuring Enablement Restriction
Habits	Repeatedly performed behaviours in stable contexts with little thought or deliberation	Incentivisation Restriction Coercion Environmental restructuring
Emotion	Actual or anticipated feelings in response to performing a behaviour	Persuasion
Biases	Systematic and unconscious tendencies to think, decide and behave in certain ways, leading to predictable deviations from a perceived standard of rationality or good judgement	Environmental restructuring Persuasion
Context	Structures and architecture in the environment that influence behaviour Shared social and cultural expectations that exist beyond any individual or single behaviour	Environmental restructuring Persuasion

Another case study: Between 2018 and 2020 we partnered with a hospital to improve their recycling in the operating rooms. During the Deep Dive we identified that staff were having difficulties accessing recycling bins due to their position in the room. Additionally, recycling bins for a specific stream (Kimguard, which is a wrapping material to keep instruments sterile) did not exist in the operating rooms. Drawing on Table 3 we can see that these barriers fall under the behavioural influence categories “Context” and “Opportunity” respectively and the intervention function “Environmental restructuring” can help address both. Taking these insights into account, together with the hospital and project funders, we designed an intervention that introduced new Kimguard recycling bins. Additionally, new bins on wheels for the existing recycling streams were introduced so that staff could easily pull recycling bins closer for better access while still having the possibility to move them “out of the way” when the situation in the operating room required it. The hospital has since reported back to us that the project, together with a lot of hard work and input from the hospital’s own enthusiastic theatre green team, helped to divert an estimated five tonnes of Kimguard from landfill over one year.⁸

Box 1: Taking intervention design further: Understanding the active ingredients of interventions

The nine intervention functions in Table 1 describe interventions at a fairly high level. But at the core of interventions sit components of change based on “first principles”. These are sometimes also called *behaviour change techniques* (BCTs; Michie et al., 2013) or “change methods” (Bartholomew Eldredge et al., 2016). You can think of these techniques as the active ingredients of an intervention. They are considered the smallest unit or

⁸ This project was part of the larger “Waste Education in Healthcare Project” delivered by the Victorian Health Building Authority, with support from Sustainability Victoria through the *Victorian Waste Education Strategy*.

“component” of an intervention that could bring about change (Michie, Johnston, & Carey, 2020).

For example, an interactive training video that teaches clinicians how to conduct escalation of care conversations with their colleagues might have an intervention function of training, as it involves imparting skills. However, if we dig beneath the surface, we can see that this training may lead to behaviour change through a number of more specific BCTs, such as (1) instructions on how to perform the behaviour, (2) behavioural practice and rehearsal, and (3) graded tasks. These may be related but distinct techniques; by distinguishing between them we have a better shot of understanding which one was responsible for changes in behaviour.

There are different classification systems of BCTs or change methods. One that we like to use is called the Behavior Change Technique Taxonomy (v1), which was again created by Michie and her team (2013; 2014; yes, they have completed a lot of influential and very helpful work in this space!). This taxonomy includes 93 different BCTs across 16 different groups, which was developed through a series of consensus exercises by behavioural experts from around the world. It provides a common language among practitioners across a wide range of contexts and disciplines to understand the specific mechanisms which lead people to change their behaviour.

These BCTs include a great level of nuance and their application might require a degree of behaviour change expertise. To facilitate the use of BCTs for intervention design, Michie and her team have created an online [Theory and Techniques Tool](#). This tool features a matrix which shows whether, based on the existing evidence and expert opinions, a BCT is linked to a “mechanism of action”, or the processes through which a BCT influences behaviour (Johnston et al., 2021). For an alternative taxonomy of behaviour change methods, see also Kok et al. (2016) and Bartholomew Eldredge et al. (2016).

BRINGING TOGETHER MULTIPLE SOURCES OF KNOWLEDGE: LITERATURE REVIEWS AND STAKEHOLDER CONSULTATION

So far, we’ve presented a “toolkit” of intervention functions and have shown how this can be mapped to two frameworks: the TDF domains and the BWA Influences on Behaviour.

Additionally, we have introduced a more advanced approach (Box 1) that maps directly to BCTs/change methods.

The key take-home messages are:

- Using a more systematic and comprehensive framework such as the intervention functions allows you to select from a range of different intervention options which go beyond the “usual suspects” of education or financial incentives.
- Designing interventions that address the identified behavioural influences of your target audience will increase the likelihood of success of your intervention. However, even the best intervention design can’t guarantee that your intervention will be successful. This is why testing and trialling your intervention is a crucial step in our BWA Method (you will be able to learn more about trials and evaluation in a later chapter).

As we cautioned earlier, it is important to draw on multiple sources of knowledge to make an informed decision about which intervention and/or technique might have a promising chance of success. In addition to the “topdown” approach of mapping categories of behavioural influences to intervention functions, you may also want to incorporate evidence gathered from “bottom-up” approaches, such as literature reviews (see Box 2) and stakeholder consultations on your specific problem. We have described both methods as part of the Exploration phase of the BWA Method, in which they can help unpack a problem. But they can also be used during the Deep Dive to gather existing knowledge on what works and to understand multiple perspectives to identify intervention options in your specific context.

As a final consideration in your intervention design, you need to determine the mode of delivery for your intervention (Michie et al., 2014). As mentioned earlier, this is the format the intervention takes. A training intervention could for example be delivered via a face-to-face workshop, an online course, or a mobile app.

Box 2: Drawing on a literature review to inform intervention design

A literature review isn’t only important for unpacking a problem during the Exploration phase of the BWA Method. Prior to your intervention design, you can also conduct a literature review to find out which interventions are likely to change your target behaviour or your audience’s behavioural influences and which interventions have failed to do so in the past. In our escalation of care case study, such a review might look for studies that have tested interventions to increase the skills and confidence of nurses, midwives or doctors having escalation of care conversations.

Some studies you are likely to find might address the details of how their intervention changed behavioural influences which in turn led to behaviour change. Others, however, might only report if an intervention led to a change in the target behaviour and not provide any details about behavioural influences. These studies will be less helpful when trying to pair an intervention with your identified behavioural influences and understanding the mechanisms through which change might occur. They will, however, still be very valuable because they give you an idea of different interventions worth trialling and those that might be less promising (i.e., interventions that proved unsuccessful in previous studies).

Another case study: One example of this is from a project in which we partnered with the charity Inclusive

Australia to improve diversity and social inclusion in the workplace. A key part of this project was to conduct a rapid review of the literature to understand whether existing diversity and inclusion interventions, practices, and policies in organisations were effective. The review revealed that there were a number of broad categories of interventions, ranging from diversity training and education to standardised job selection procedures. While the evidence was relatively stronger for some interventions such as workplace accommodations, it was weaker for others, such as those promoting visibility of marginalised groups. The review also revealed what sort of specific outcomes could be expected for a given intervention. For instance, diversity training was associated primarily with changes in knowledge and awareness, but less so with behaviour change. These findings would therefore play an important role in guiding decision-making when it comes to designing our own intervention for workplace diversity and inclusion.

Finally, it is important to keep in mind the limitations of this approach. Just because an intervention hasn't worked in another setting, population, or country doesn't mean that it won't work in your current context. Again, this approach is intended to give your intervention the greatest likelihood of success based on existing evidence, but you may also need to consider multiple sources of information and expert opinion from stakeholders. This is why the co-design approach is so crucial.

INTERVENTION PRIORITISATION

By now you may have a list of intervention options, informed by your mapping of behavioural influences to intervention functions or BCTs/change methods, as well as any stakeholder advice and review of the literature. The next step is to prioritise which of these to develop further and implement. Much like we can prioritise target behaviours based on criteria like impact, likelihood, and current adoption (see Chapter 5), we can prioritise interventions on a number of criteria following a similar process. Prioritisation essentially involves going from a long list to a short list. In this section, we briefly outline key steps and possible criteria that could help you with the prioritisation process to select your intervention. At its core, the prioritisation process requires the essential steps of: (1) selecting prioritisation criteria that are relevant for your problem and your project and (2) applying a prioritisation method to compare and evaluate the potential interventions against each other. [Selecting prioritisation criteria](#)

There are a number of possible criteria you could use to reduce your “long list” of intervention options to a “short list”, in order to reach your selected intervention. For instance, we have already introduced concepts such as *impact* and *likelihood* (sometimes in the context of an intervention referred to as “reach”), but you may also like to consider the *potential side-effects or unintended consequences* of each intervention (e.g., an intervention to reduce food waste could encourage the use of take-away containers to take home leftover food when eating out, but using these containers could lead to more waste) or its potential *scalability* (scaling and knowledge translation will be covered in a later chapter). In addition, you may wish to consider the *feasibility of the interventions*, regarding whether they might be more or less acceptable to the community, their political sensitivity, and whether they align with your own organisation's or your partner organisation's remit.

As an illustration of feasibility considerations, after reviewing the existing intervention options, you might feel like some don't fall under your organisation's remit in the system that surrounds your target audience and target behaviour. For example, you might have no influence on any existing regulation. Such a situation can limit you to certain intervention functions which fall under your area of influence. However, in instances where you identify key intervention types outside of your competences, it might be worthwhile exploring who you can partner with so these interventions become part of the mix you (through your project partners) can apply to tackle your problem.

Other intervention prioritisation criteria might include the *ease with which the intervention could be implemented* and the *extent to which you could monitor and evaluate the intervention*. Finally, there may be a number of factors that restrict or limit which intervention could be implemented (*constraints*), such as funding or resourcing and project timelines.

Your choice of prioritisation criteria can also be guided by criteria other behavioural experts have developed, such as APEASE (Affordability, Practicability, Effectiveness and costeffectiveness, Acceptability, Side-effects and safety, Equity), used by Michie et al. (2014).

We recommend exploring which criteria are of greatest relevance and importance for your project together with any partner organisation(s) and/or key stakeholders, while keeping things practical (e.g., a "top five")—you may have already explored key priorities and constraints at an earlier stage of the BWA Method, which you could draw on during the intervention design phase. [Applying a prioritisation method](#)

Assuming you have identified your prioritisation criteria, you will now need some way of applying them to help you decide which intervention(s) make it to the "short list", and ultimately to select your intervention. There are, of course, different degrees to which a prioritisation method may be applied, ranging from a relatively informal discussion between your team and your partner organisation to a formal prioritisation workshop with key stakeholders.

How to apply a formal prioritisation method using multi-criteria analysis in the context of selecting target behaviours was discussed in Chapter 5. In much the same way, a multicriteria approach to selecting interventions would essentially involve assigning a "score" for each intervention option across each criterion. Formal tools to assist with scoring and comparing your interventions based on your criteria include a prioritisation matrix (e.g., see Chapter 5) or an intervention report card (see Figure 1).

Again, what information or evidence is used to inform criterion scores could range from project team and/or expert opinions (which can be very basic, such as the traffic light system in the report card) to assigning scores based on existing evidence or gathered data. For example, to gauge how feasible your intervention options might be from the perspective of their acceptability, you could measure perceptions of the interventions via a community survey.

After scoring each intervention option across your criteria, the interventions can then be compared by ranking them according to their total score (e.g., with higher total scores reflecting interventions that could be prioritised) (for a thorough discussion about using multicriteria decision analysis, in the context of health interventions, see Baltussen &

Niessen, 2006). Once you have compared your interventions, and decided which one to develop, you can proceed to the fun part: designing your final product.

Target behaviour:	
Intervention:	
Respond using the following:	Rating
<input checked="" type="radio"/> Yes	
<input type="radio"/> Maybe	
<input type="radio"/> No	
<input type="radio"/> Can't say	
Reach: Will a large part of your audience be exposed to the intervention?	
Impact: Assuming they are, will the intervention make a big difference to the target outcome?	
Time frames: Can the intervention be run in 7 months?	
Monitoring and evaluation: Can you directly measure the behaviour targeted by the intervention?	
Other comments?	

Figure 1. An example of an intervention report card used for prioritisation. This can be a useful format if you wish to retain brief, summarised records of ideas from a prioritisation exercise for future reference.

PUTTING EVERYTHING TOGETHER: A GUIDE TO INTERVENTION CODESIGN

We now understand all the steps required to design an intervention. But how do you actually turn this knowledge into a tangible intervention or intervention package, for example a training program, new legislation or an awareness raising campaign? There are a variety of approaches for putting everything together in order to design an intervention. One recommended way to do this is through an intervention co-design process, in which different stakeholders are brought together to develop a shared understanding of the behavioural problem and to design an intervention⁹. Below, we provide a brief guide on how to carry out an intervention co-design, followed by a list of common pitfalls that we frequently encounter during this process (Box 3).

What is an intervention co-design?

This is a participatory process for designing a behaviour change intervention which brings together various stakeholders—target audiences and beneficiaries, governments, local authorities, field partners, local staff, collaborators, program managers, and behaviour change experts.

In a co-design, all stakeholders are designers of the behavioural intervention, and are involved in key stages of co-defining the problem, understanding potential solutions, codevelopment of ideas, and co-implementation of solutions (Robertson & Simonsen, 2013). Although there is no single approach to co-design, it is characterised by the sharing of decision-making power so that equal value is given to the expertise provided by each

⁹ This is our recommended approach in most circumstances, but note that there are other approaches for intervention design, such as a hackathon, in which a team of behavioural experts work through an intensive “design-sprint” to develop recommendations for an intervention.

stakeholder (i.e., scientific expertise, practical expertise, professional expertise, expertise of lived experience) (Burkett, 2012).

This helps ensure that the intervention:

- draws from multiple sources of expertise and the collective creativity of stakeholders,
- is suited to the local context,
- aligns the needs and interests of multiple parties, and
- has local ownership and is supported and adopted by key stakeholders, delivery staff, and the target audience (Jessup, Osborne, Buchbinder, & Beauchamp, 2018).

What is involved in an intervention co-design?

An intervention co-design consists of three broad stages: co-definition, co-development, and co-implementation (Carvalho et al., 2017).

Step 1: Co-definition

Developing a mutual understanding and getting on the same page: If your issue at hand involves some level of complexity, we recommend beginning the co-design process with an activity that creates a common understanding and vocabulary of your issue. The purpose of this activity is to emphasise the role of each stakeholder as well as ensuring a smoother communication process for the remainder of the co-design.

Such an activity could consist of a mapping exercise (see also Chapter 2 on systems mapping). To illustrate this, in the project we mentioned earlier to improve waste separation in hospitals, we ran a co-design workshop with different hospital staff members from different teams. We quickly realised how complex this space is and started the workshop with a mapping exercise that outlined the waste flow. This visual outline proved vastly beneficial, as throughout the workshop participants could draw on their waste map to illustrate what they were referring to. It is also important to begin by building consensus around a definition of the target behaviour (consisting of all its AATCT components; see Chapter 5). You may already have such a definition from the beginning of the Deep Dive phase.

Providing background information: Once you have established this common ground on your problem and target behaviour, the next step is to provide all participants with the necessary information for designing the intervention. This includes all the insights gathered from the Deep Dive and the Exploration phase of your project, such as any output from surveys, interviews and observation (Chapter 7), as well as literature and practice reviews (Chapters 1 and 4). This is also a good opportunity to direct participants to behavioural

science principles, concepts, and techniques given the behavioural nature of your intervention.

Step 2: Co-development

Ideating for intervention design: There are a number of ways of running the design phase of co-development, which will depend on your group, project, available time, and your own preferences as facilitator. One common approach is to begin with a general ideation process in smaller groups, followed by a larger group discussion where participants can bounce ideas off one another and come up with modifications or improvements to existing interventions. The key requirement is to ensure that your stakeholders are well-represented across these groups.

Throughout this process, you will need to have your toolkit of intervention functions for reference. This is the stage where a behavioural expert plays a critical role in ensuring that behavioural influences are matched to intervention functions, and that all appropriate intervention functions have been considered.

Prioritising your intervention: You will now have a number of ideas for different interventions. Using the criteria developed during the last section, you can begin to prioritise your interventions with the advice of your stakeholders. One quick approach we use is to have all participants complete a report card (Figure 1). This is also a good opportunity to check with stakeholders (especially those who will be assisting with running the intervention on the ground) whether all relevant feasibility criteria have been included.

Step 3: Co-implementation

Planning for the selected interventions: This stage aims to reach an agreement from all stakeholders on the next steps and who will be responsible for them. This ensures buy-in from stakeholders and provides you with a contact for following up remaining questions and detailed planning activities. Questions to consider include:

- Which tasks need to be done, step by step?
- Who is leading the completion of the task?
- Who is supporting the completion of the task and how?
- By which date will the task be completed?
- Are participants currently making any assumptions regarding the implementation (e.g., external support from another stakeholder)?

Another case study: In Chapter 7, we introduced you to a project that explored the drivers and barriers associated with transferring patients following an acute spinal cord injury (SCI) from an indwelling catheter (IDC) to an intermittent catheter (IC), as this has been shown to reduce the incidence of urinary tract infections. Codesign with a range of stakeholders, including clinicians and patients, was an important part of the process, from identifying potential behaviour change techniques to the final intervention. For example, one of the drivers of clinician behaviour was beliefs about the capability of patients following an acute SCI to both use ICs and make an informed decision and therefore consent to this change in

bladder management. At that time, the only form of patient information on optimal bladder management was contained in a large folder that had very technical and lengthy information on the pros and cons of each method. One of the main problems was some patients could not even access the folder without support, as they did not have the ability to hold the folder or easily turn pages as a result of the SCI. To ensure a new website was both informative and attractive to people with a newly acquired SCI, we worked with the Australian Quadriplegic Association and people living with SCI to co-design the new resource. In addition to the content, this also ensured that it was accessible via touch-based screens that would enable people with limited hand function to easily access the information. The new website (www.mybladdermylife.com) is now accessible and iPads are available at the hospital to ensure everyone can access the information. The process of co-design also supported the implementation of the interventions, as there was greater awareness of the project aim and subsequent engagement with, and promotion of, desired behaviours.

Box 3: Pitfalls to avoid during intervention design

As we have discussed so far in this chapter, there are clear steps that we can take (and tools that we can use) to help design better interventions. However, based on our experiences, there are also several pitfalls that we should be aware of and make plans to avoid.

Defaulting to the same types of interventions: “*This is what we’ve always done!*” is a remark we often hear when working with stakeholders and partners. On the one hand, this approach might mean that your team has expertise in this area, but it also runs the risk of relying on the same interventions while missing out on other more effective and novel solutions. This can be remedied by systematically mapping potential interventions to your behavioural influences, which ensures that you don’t miss these blind spots. Be mindful of stakeholders who come with strong opinions or preconceived ideas about what will work. The intervention design stage is an opportunity to broaden and challenge thinking, especially beyond what has been tried in the past.

Adding rather than subtracting: One phenomenon that we often notice when designing interventions is the tendency to *add*—for example, adding processes, documents, tasks and programs—rather than considering how we might subtract or remove to achieve a better outcome. In hospitals where errors or incidents occur, investigations into why these events happen often lead to another procedure or guideline being added to the current process (in one circumstance, this resulted in over 30 documents added!). An alternative approach is to determine what processes could be removed, which would, for example, reduce cognitive load and streamline activities. The book Subtract, by Leidy Klotz, examines how we can build subtraction in our thinking to design better behavioural solutions.

Confusing the mode of delivery with the intervention itself: One common misconception that we come across is the tendency to confuse the mode of delivery (such as a mobile application) with the type of intervention, such as training (which could be applied through a number of different modes of delivery). Again, this is where systematic

frameworks of behaviour change interventions from the start of the chapter come in handy, as they are accompanied by clear definitions that delineate between similar concepts.

Failing to consider implementation in the intervention design: Often we find that issues of implementation (or the actual “putting into practice” of your intervention) are considered too late (e.g., when the intervention is being rolled out). However, the two are linked and you need to be thinking about implementation (and related issues such as scalability) when designing your intervention. If you are applying an iterative design process, chances are you will already be building on learnings derived from early implementation during your intervention design. You’ll also need to ensure that you have appropriate expertise (e.g., including communication specialists) for your intervention type. It is difficult to “add on” this expertise after an intervention has been determined, so it is best to include all relevant perspectives when developing and prioritising intervention options. The later chapters of this book will cover implementation in more detail.

Back to our escalation of care case study. We began our intervention design by reviewing the evidence gathered from our Deep Dive on the drivers and barriers around escalation of care conversations. We convened a stakeholder dialogue to deliberate on the gathered evidence and intervention options. The stakeholder dialogue allowed us to consider and prioritise appropriate intervention options, having identified education and training as addressing the behavioural drivers, and considered prioritisation criteria such as feasibility in the hospital environment and scalability.

After all of these steps, we landed on a final intervention that consisted of an interactive choose-your-own-adventure training video to address the knowledge and skills gaps identified during the Deep Dive. This interactive training video was designed to help clinicians gain experience in navigating some of the key behavioural barriers to timely escalation of care, such as fear of hierarchy and pushback to escalation of care.

The video took participants through five scenarios during escalation of care conversations between clinicians. Each scenario involved a particular challenge and a “knowledge lesson” at the end.

But did the intervention lead to improvements in clinician experience and skills, as well as changes in the number of escalation of care conversations as we had intended? A field trial in two public Victorian hospitals observed positive effects in increased clinician confidence to have conversations and reduced their reluctance to escalate care. The number of challenges reported during escalation of care conversations was also reduced.

CONCLUSION

This chapter has provided an overview of the steps involved when making sense of the insights gathered during the Deep Dive and putting these together to design an intervention. We have outlined some key considerations, such as the different intervention functions available to you, mapping behavioural influences to intervention functions, BCTs or change methods, prioritising intervention options, and the importance of co-design. Once these steps are complete, you will be ready to carry out your behaviour change intervention and test its effectiveness, which will be covered in the final chapters of this book.

But sometimes, you may not be able to invest the resources to complete a thorough Deep Dive and understand the influences on your target behaviour. In other instances, the target behaviour may be relatively simple and you already have a great deal of knowledge so that a Deep Dive would be unnecessary. In these cases, what could you do? It turns out that there are several general models of behaviour change that you have at your disposal, which can guide intervention design in the absence of a targeted Deep Dive. We will go through these in the next chapter.

REFERENCES

- Baltussen, R., & Niessen, L. (2006). Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost Effectiveness and Resource Allocation*, 4,14. doi: 10.1186/1478-7547-4-14
- Bartholomew Eldredge, L. K., Markham, C. M., Ruiter, R. A. C., Fernández, M. E., Kok, G., & Parcel, G. S. (2016). *Planning Health Promotion Programs: An Intervention Mapping Approach*. San Francisco, CA: Jossey-Bass.
- Burkett, I. (2012). *An Introduction to Co-Design*. Sydney: Knode.
- Cane, J., O'Connor, D., & Michie, S. (2012). Validation of the theoretical domains framework for use in behaviour change and implementation research. *Implementation Science*, 7, 37. doi:10.1186/1748-5908-7-37
- Carvalho, F., Jun, G. T., & Mitchell, V. (2017). Participatory design for behaviour change: An integrative approach to healthcare quality improvement. In *IASDR 2017 Proceedings, 7th International Congress of the International Association of Societies of Design Research*.
- Darnton, A. (2008). *GSR Behaviour Change Knowledge Review. Reference Report: An overview of behaviour change models and their uses*. London: HMT Publishing Unit.
- Dolan, P., Hallsworth, M., Halpern, D., King, D., Metcalfe, R., & Vlaev, I. (2012). Influencing behaviour: The MINDSPACE way. *Journal of Economic Psychology*, 33(1), 264-277. doi:10.1016/j.jeop.2011.10.009
- Gould, D.J., Moralejo, D., Drey, N., Chudleigh, J.H. & Taljaard, M. (2017). Interventions to improve hand hygiene compliance in patient care. *Cochrane Database of Systematic Reviews* 2017, 9, CD005186.
- Jessup, R. L., Osborne, R. H., Buchbinder, R., & Beauchamp, A. (2018). Using co-design to develop interventions to address health literacy needs in a hospitalised population. *BMC Health Services Research*, 18, 989.
- Johnston, M., Carey, R. N., Connell Bohlen, L. E., Johnston, D. W., Rothman, A. J., de Bruin, M., . . . Michie, S. (2021). Development of an online tool for linking behavior change techniques and mechanisms of action based on triangulation of findings from literature synthesis and expert consensus. *Translational Behavioral Medicine*, 11(5), 1049-1065. doi:10.1093/tbm/iba050
- Kok, G., Gottlieb, N. H., Peters, G.-J. Y., Mullen, P. D., Parcel, G. S., Ruiter, R. A. C., . . . Bartholomew Eldredge, L. K. (2016). A taxonomy of behaviour change methods: An Intervention Mapping approach. *Health Psychology Review*, 10(3), 297-312. doi:10.1080/17437199.2015.1077155

- Michie, S., Atkins, L., & West, R. (2014). *The Behaviour Change Wheel: A Guide to Designing Interventions*. Great Britain: Silverback Publishing.
- Michie, S., Johnston, M., & Carey, R. (2020). Behavior change techniques. In M. D. Gellman (Ed.), *Encyclopedia of Behavioral Medicine* (pp. 206-213). Cham, Switzerland: Springer Nature Switzerland.
- Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., . . . Wood, C. E. (2013). The Behavior Change Technique Taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine*, 46(1), 81-95. doi:10.1007/s12160-013-9486-6
- Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, 6, 42. doi:10.1186/1748-5908-6-42
- Robertson, T., & Simonsen, J. (2013). Participatory design: An introduction. In *Routledge International Handbook of Participatory Design* (pp. 1–17). New York, NY: Routledge.
- World Health Organization, (2009). *WHO Guidelines on Hand Hygiene in Health Care: First Global Patient Safety Challenge: Clean Care is Safer Care*. Geneva, Switzerland: WHO Press.



MONASH
University

MONASH
SUSTAINABLE
DEVELOPMENT
INSTITUTE

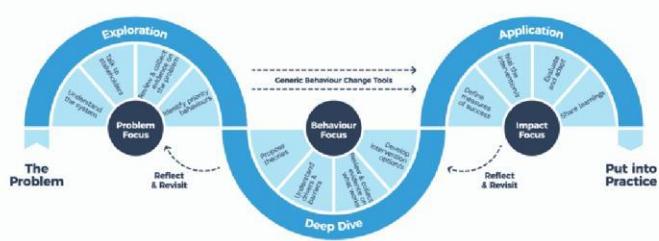
BehaviourWorks
AUSTRALIA

THE METHOD

CHAPTER 9: Using generic behaviour change tools when time and resources are scarce

By Liam Smith, Peter Slattery, Brea Kunstler and Jenni Downes

www.behaviourworksaustralia.org



INTRODUCTION

One of our partners, VicRoads, was concerned about the number of people who were driving on the road when they probably shouldn't be. Typically, potentially unfit drivers get reported to VicRoads by a family member, a friend or the police. Once this happens VicRoads writes to them and asks the individual to go and see their doctor for a medical check to ensure they are able to continue to drive. After the visit to the doctor, potentially unfit drivers need to submit paperwork to VicRoads within four weeks stating that they are able to drive. Doing so prevents cancellation of their licence and allows them to continue to drive. The problem was that many people weren't submitting their paperwork. Half in fact. This was worrying for a few reasons. Most notably, by continuing to drive, they may be putting others at risk but also because in the event of an accident, they'd be uninsured.

Usually in situations like this, we'd advocate undertaking a Deep Dive (see Chapters 6 to 8) to understand the reasons why drivers were not submitting their paperwork. However, in this situation, we had limited time and resources, and accessing the target audience (i.e., potentially unfit drivers) to understand the reasons behind nonsubmission was going to be difficult. So, the question we asked ourselves was how could we help address the problem in a narrow time frame without first doing a Deep Dive with our audience?

During the Exploration phase (Chapters 1 to 5), we aim to understand the problem and identify and prioritise behaviours to change. Once we have identified priority behaviours, a key decision is to decide whether a Deep Dive (i.e., a more detailed investigation of the behaviour and the target audience/s) is desired or needed. Completing a Deep Dive phase gives us a fuller understanding of the drivers and barriers to particular behaviours, and based on this understanding, gives us guidance to design tailored solutions. In general, we prefer this approach because tailored interventions that are based on a deeper investigation of the problem and behaviour tend to be more effective than those that are untailored (Gans et al., 2009).

However, there are occasions when we need to take a more streamlined approach. In these cases, we might choose to use tools that we know might still be effective. At BehaviourWorks Australia (BWA), we use the term 'generic behaviour change tools' to refer to 'off the shelf' approaches which have been shown to work in a variety of circumstances (albeit rarely, or not as well, as tailored approaches). There are a wide range of generic behaviour change tools and using them is better than guesswork when time and resources are scarce.

In this chapter, we outline three 'generic behaviour change toolkits' that we frequently use in our work, as well as the circumstances when we might choose to use them. These toolkits are BWA's own INSPIRE framework, the Behavioural Insights Team's EAST framework, and Robert Cialdini's Principles of Persuasion.

WHEN SHOULD I USE GENERIC TOOLS AND WHEN SHOULD I DEEP DIVE?

Whether to start with a Deep Dive or to understand how beneficial Deep Dive (vs using generic tools) will be, the following questions are worth considering:

Is the behaviour a key focus for you or your organisation? For important behaviours, such as those that align with your strategic priorities, we would always recommend a Deep Dive. For example, if your organisation's primary focus is on birth control in teenage girls, or vaccination in indigenous communities, then a deep Dive is highly recommended for behaviours that align to these focus areas, as it will increase the chances of designing effective behaviour change interventions. However, in situations where there are many target behaviours, more generic behaviour change tools may be appropriate.

Is the target audience large? For large audiences, generic tools have been shown to achieve small but collectively impactful changes. Consider, for example, communication that encourages all citizens to get vaccinated. Communication may only increase the number of people doing the behaviour by a few percent, but this may translate to thousands of people and may even be enough to push the community to a level of herd immunity.

Do you or your organisation want to know why the intervention worked or failed? The Deep Dive will tell us what drivers and barriers need to be targeted to increase the chance of behaviour change. Because not all interventions work, having this 'theory of change' can be very helpful to evaluate the intervention's impact, including whether the targeted drivers and barriers changed, even if the behaviour didn't. This allows us to understand where our theory fell down and plan different interventions in the future. To explain with an example, we may believe that telling people that trains are cleaned regularly will increase patronage. If we implement a cleaning regime and a communication intervention saying this, we can first evaluate if people became more aware that trains were cleaned, as well as whether we see increased patronage in those that had this awareness. Without this theory, we're only guessing whether our cleaning regime was noticed or our message was received.

Is there time and budget to investigate? Often time or budget constraints prevent the opportunity to undertake Deep Dives to understand drivers and barriers to desired behaviours. In these cases, generic behaviour change tools may represent a way forward.

Is the behaviour complex? Simple, momentary, one-off or infrequent behaviours can lend themselves to simple tools. Purchasing decisions in a supermarket, filling out a form correctly and not parking illegally are all behaviours that can be 'nudged' using generic behaviour change tools. More complex behaviours, such as major purchases (e.g., homes, solar panels, cars), swapping travel modes and choosing career options, are more detailed and mindful behaviours which are better understood using Deep Dive tools.

As you can see, there are a number of considerations when deciding to Deep Dive (or not). Depending on your answers to the above questions, you might find that generic tools suit your circumstances and priorities better. In the remainder of this chapter, we outline three toolkits that we frequently use. There is no exact science for the choice of any of these - all contain tools that can be helpful in most situations.

THE INSPIRE FRAMEWORK

The INSPIRE framework (Faulkner et al. 2019) was developed by BWA researchers as a simple mnemonic to describe seven powerful techniques for improving compliance requests

through written communication. Each letter represents a tool that can be used to change behaviour, and have been shown to do so in different contexts. These letters stand for:

- **Implementation intentions**
- **Norms**
- **Salience**
- **Procedural justice**
- **Incentives**
- **Reputation and credibility, and**
- **Ease**

Implementation intentions are a tool that can be used to overcome the gap between people's intentions and their actual behaviour. This gap has been a longstanding challenge for behaviour change practitioners. Implementation intentions work by encouraging the audience to identify plans to implement a behaviour (or overcome obstacles that may prevent it) and have been shown to be effective in several contexts such as getting vaccinated (Milkman, Beshears, Choi, Laibson, & Madrian, 2011), medical screening (Milkman, Beshears, Choi, Laibson, & Madrian, 2013), and exercising (Milne, Orbell, & Sheeran, 2002). Typically, such plans take the form of writing down a time and place where you will do the behaviour and / or writing "if-then" plans to consider where and when the behaviour will occur and ways to navigate around obstacles.

Norms are our perceptions of others' performance and opinions about behaviour. The most common application of norms is to inform the target audience that desirable behaviours are common, using phrases such as "nine out of ten taxpayers pay their tax on time" or "the vast majority of Australians intend to get vaccinated". The notion that these behaviours are already being undertaken makes people attracted to behave similarly, often in subconscious ways. Norms can also work in negative ways because they can also tell people that a lot of people are doing an undesired behaviour. For example, phrases such as "there's an obesity epidemic" and "thousands of people aren't following COVID restrictions" can attract people to follow suit, again often in unconscious ways (Schultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007). Norms can also refer to our perception about what others would think of our behaviour. If we feel others would approve, we're more likely to comply. Interventions based on this type of norm typically include approvals for past behaviour (e.g., a tick or a smiley face) and come from people who are seen as peers.

Salience refers to the use of stimuli that demands attention, such as personalisation, colour and images, to attract the audience to pertinent information needed to change the behaviour. Researchers Pippa Scott and Phil Edwards reviewed many approaches to encourage people to return questionnaires and found that personalised letters and using handwriting (rather than typing) to address letters or writing the letters themselves, were effective (Scott & Edwards, 2006). Colours have been used to make it clear that bills are overdue (Service et al., 2014) and to encourage online vehicle registration (Castelo, Hardy, House, Mazar, Tsai & Zhao, 2015).

Procedural justice involves the communicator being honest, transparent, unbiased and respectfully treating those receiving communication. Procedural justice techniques typically take the form of explaining why individuals have been contacted and why punishments are necessary (but don't blame the individual) and work best in regulation contexts. They have

been used to encourage tax payments (Wenzel, 2006) and to encourage potentially unfit drivers to get medical check-ups (Faulkner, Jorgensen, Sampson & Ghafoori, 2018) (more on this later).

Incentives involve promising items or gestures that the target audience value to persuade them to change their behaviour. Such valued items might include cash or entry into lotteries, or non-tangible items such as positive recognition and appreciation. There are numerous examples of incentives being used to encourage behaviour of a whole range of contexts such as returning questionnaires (David & Ware, 2014), encouraging exercise (Mantzari et al., 2015) and applying for work (Dal Bó, Finan, & Rossi, 2013). The summary of this research suggests that, in most cases, the use of incentives is effective in encouraging behaviour change.

Reputation and credibility involves using credible and trustworthy sources to deliver information to the audience who is being asked to change behaviours. It is important to ensure that the deliverer of information is truly considered credible and trustworthy by the audience, and is not just expected to be perceived as credible and trustworthy. Much research has looked at features of the source that make the message more impactful. In general, this research shows that people with particular jobs (such as doctors, nurses and teachers) (Brenan, 2017), people who are perceived as similar and have relevant experience (Phua, 2016), people who are seen as authorities, and people we like (Cialdini, 2009), are more credible. Finding people who meet as many of these criteria as possible to deliver messages will increase the likelihood of compliance.

Ease involves first making the behaviour as easy as possible by removing as many barriers as possible. Having a single direct contact number for enquiries or making services available locally are examples of removing barriers. For written communication, ensuring instructions are clear, easy to read and that the requested behaviour is straightforward to perform is important. Making language simple to understand, altering text layout (chunking) and including greater text spacing have also been shown to increase responses to written communication (Faulkner, Borg, Bragge, Curtis & Ghafoori et al., 2019).

We have sought to summarise the tools of INSPIRE here in a few sentences. Appendix 1 of this chapter includes a table of the included techniques and ways to apply them and further reading can be found in a paper published by BWA researchers on INSPIRE in the journal *Public Administration Review* (Faulkner, Borg, Bragge, Curtis & Ghafoori et al., 2019).

Returning to our opening case study of potentially unfit drivers who need to submit their paperwork, we used elements of INSPIRE because we were limited in the time we had to Deep Dive. Accessing drivers to talk to was also going to be difficult and the conversations might have been awkward. So, we opted to use elements from INSPIRE to write a new letter for VicRoads to send out.

An example of the original letter with the medical report request is provided in Figure 1, while an example of the new version using INSPIRE is provided in Figure 2. Some of the specific changes we made were to simplify the language and chunk it under headings so it was easier to read and understand (ease). We were also conscious that receiving a letter from VicRoads saying a driver may be unfit to drive may be quite confronting to someone who

values driving. We therefore used procedural justice approaches to make the letter empathic to the situation ('we understand that receiving this letter can be hard for some people'), to make it clear that the driver was being treated fairly, and the reasons why suspensions are imposed ('to make the roads safer for everyone'). These relatively simple changes, among the use of other INSPIRE techniques, were implemented and tested against the 'usual' letter that VicRoads sent out.

Figure 1: Example of the original letter with the medical report request

17 December 2015
Our Ref: 40000475

Medical Review
PO Box 2504
Kew VIC 3101
medicalreview@roads.vic.gov.au

Mr Medical Review Testing
L 1 60 Denmark St
KEW VIC 3101

Phone: (03) 8391 3226
Web: vicroads.vic.gov.au

Dear Mr Testing

REQUEST FOR MEDICAL REPORT

VicRoads is responsible for driver licensing in Victoria. An important part of this role is to ensure that people who operate vehicles are competent drivers and are medically fit to hold a [[driver licence / learner permit]].

I have recently become aware that your health may be affecting your ability to drive safely. For this reason, you need to arrange for a doctor who is familiar with your medical history to complete the enclosed medical report. Depending on the content of this report, you may also be asked to take a driving test.

When making the appointment, you should tell the receptionist that you are coming in to have this report completed, as the examination may take longer than a standard consultation. Please take all the documents in this envelope with you to your appointment, and write your name, address and [[driver licence / learner permit]] number on the report and sign it before giving it to your doctor.

Your doctor needs to return the report to VicRoads in the enclosed envelope within four weeks from the date of this letter. Otherwise your [[driver licence / learner permit]] may be suspended or cancelled in accordance with regulation 78(1) of the Roads Safety (Drivers) Regulations 2009. Further details, including important information for probationary drivers, are provided on that back of this letter.

Please see the enclosed brochure for more information. If you have any question, please contact Medical Review on the above phone number.

Yours sincerely



MANAGER MEDICAL AND DRIVER REVIEW

Figure 2: Example of the revised letter based on INSPIRE

Date: 15 December 2015
Our Ref: [[Reference number]]

VicRoads Medical Review
PO Box 2504
Kew VIC 3101
Tel: (03) 9854 2407
Fax: (03) 9854 2307
medicalreview@roads.vic.gov.au

Dear [[John Citizen]],

Request for a medical report

At VicRoads, we often receive reports from police, doctors, and community members expressing concerns about individuals' medical fitness to drive. Unfortunately, we have recently received a report about your medical fitness to drive.

Why are we sending you this letter?

When we receive these reports we are legally obliged to investigate, even if you have a perfect driving record. Everyone has a role to play in keeping our roads safe. In this case, our role is to ensure that all drivers in Victoria meet national medical standards for licensing.

What you need to do

The first step of this medical review is for you to provide us with medical report from a doctor.

When you need to do it

Please ensure that the enclosed report form is completed and returned to us by [[dd Month, Year]].

What happens after you send the report?

After we receive your report, we will write to you again to either inform you of the outcome of your review, or request additional information. Sometimes the initial report identifies issues that require more detailed examination to determine if they are affecting your ability to drive. This means that we may need to ask you to attend further specialist appointments or practical driving assessments in the future.

What happens if you don't provide the report in the time requested?

If we don't receive the completed report form by the due date above, we regret to inform you that your [[driver licence/learner permit]] will be suspended from [[dd Month, Year]].

Why do we impose suspensions?

Without medical reviews, our society would face heightened risk of serious injuries and deaths on our roads. Most people who receive this letter willingly supply the report in the required timeframe.

However, a suspension system needs to be in place for rare cases where a willingness to comply is missing.

We do not want to make things more difficult for you

We understand that receiving this letter can be hard for some people, and there may be some reason why this is a particularly difficult time for you. However, we trust you will understand the need for the medical review process, and thank you in advance for playing your part to ensure that you and all others are safe on Victoria's roads.

For further information about the VicRoads Medical Review process, please see the enclosed information sheet or visit www.vicroads.vic.gov.au.

Yours sincerely,

[NAME]
Manager Medical and Driver Review

Your medical report is due to VicRoads by:

<insert date>

If VicRoads have not received your completed report by this date, your [[driver licence/learner permit]] will be suspended on:

<insert date>

EAST

The EAST framework was developed by the UK's Behavioural Insights Team as a simple and memorable mnemonic for policy makers and practitioners (Service et al., 2014). It argues that promoting behaviour change should involve making the desired behaviour Easy, Attractive, Social and Timely. A short summary of the principles of EAST are presented here, but we encourage further reading of the framework (see https://www.behaviouralinsights.co.uk/wp-content/uploads/2015/07/BIT-PublicationEAST_FA_WEB.pdf).

- **EASY:** To make desired behaviours easy you can, for example, change the default (e.g., opt-out organ donation), or find ways to reduce the effort (or perceived effort) of doing the behaviour.
- **ATTRACTIVE:** Behaviours can be made more *attractive* by drawing attention to them, and providing incentives for performance and sanctions for non-performance.
- **SOCIAL:** Social factors can be leveraged by making the behaviour appear to be widely performed and asking for commitments to undertake behaviours via pledges and other public statements.
- **TIMELY:** Being *timely* involves delivering information when receivers are most receptive, providing immediate consequences for behaviour when possible and helping people to make plans to address barriers to performing behaviours.

Box 1: Applying EAST to changing patient behaviour

One of our authors, Brea Kunstler, is also a practicing physiotherapist. She sees many clients and they all have one thing in common - the drive to change their behaviour (e.g., return to exercise) to achieve a goal (e.g., improve their general health).

This led Brea to complete her PhD in discovering the ways physiotherapists change client behaviour. She discovered that physiotherapists don't feel like they have the qualifications (e.g., a relevant degree) or opportunities (e.g., time within consultations) to change client behaviours (e.g., increasing exercise levels). Importantly though, she discovered that physiotherapists use several different strategies to change client behaviour all the time, without even knowing it!

Brea was asked by the Australian Physiotherapy Association to write a piece on how physiotherapists can change behaviour in a busy clinical environment using simple strategies they might already be familiar with. She jumped at the chance to share the EAST Framework with her clinical peers. You can read this piece here (see <https://www.behaviourworksaustralia.org/wp-content/uploads/2019/05/Behaviouraltechniques-to-supportphysiotherapists.pdf>)

Here's a quick summary of how physiotherapists can use the EAST framework to change client behaviour (e.g., returning to exercise)

People seek assistance from physiotherapists for various concerns. For example, returning to exercise after having a baby. Here we provide an example of how a physiotherapist (Mark) can use the EAST framework to support their client, new mum Sarah, to return to exercise.

Make exercise easy and attractive for Sarah

Mark could encourage Sarah to walk to the shops with the pram (*easy*) to collect her groceries instead of taking the car. Mark should mention to Sarah how nice it would be to get some fresh air and listen to the sounds of the birds chirping as she walks, and how the gentle movement of the pram might be enough to settle her baby to sleep (*attractive*).

Make exercise social and timely for Sarah

Mark should mention to Sarah that she is one of many parents who walk with their baby to squeeze exercise into their day. In fact, Mark did this when his daughter was a newborn. This might make her feel like she is similar to other new parents (*social*). This strategy also helps Sarah contextualise walking for commuting or to help the baby settle as a behaviour common to new parents who might struggle to find time to exercise by themselves (*timely*).

Although applying this generic tool to Sarah might work to increase her exercise levels, it is important to keep in mind that all humans have their own unique drivers and barriers to behaviour. What can Mark suggest that Sarah does if the weather suddenly turns and a walk outdoors is no longer easy or attractive? This demonstrates the importance of using a Deep Dive approach if generic tools don't work as intended.

CIALDINI'S PRINCIPLES OF PERSUASION

A prominent researcher in behavioural science since the 1980s, Robert Cialdini led the way in understanding a number of behaviour change tools that have informed toolkits such as INSPIRE and EAST. Indeed, his book "Yes", published with co-authors Noah Goldstein and Steve Martin, outlines 50 generic tools that can be used to persuade people to say yes to requests (Goldstein, Martin, & Cialdini, 2010). However, his most well recognised work is "Influence: Science and Practice" which outlines six principles of persuasion (PoP), each of which relate to innate human characteristics which can be leveraged to influence behaviour (Cialdini, 2009). These principles are:

- **Scarcity:** Scarce resources are commonly valued more highly than plentiful ones. Making things or opportunities seem scarce or in limited supply can motivate audiences to behave in ways that ensure they don't miss out.
- **Reciprocity:** People generally feel obliged to repay a favour. The key to reciprocity is to give first before asking for a favour in return. Examples include writing a blog, or writing a free book (and releasing a chapter each month!).
- **Social Proof:** People are usually swayed by the actions of those around them. We tend to follow herds because there's safety in numbers.
- **Liking:** Someone who is liked by their target audience has a better chance of persuading them. We tend to like people that share similar interests or have something in common. But we are also superficial and like people who are perceived to be attractive.
- **Authority:** Authority figures can often persuade people to do things they wouldn't do otherwise. People with titles (e.g., Sir, Doctor or Professor), taller people and even people in lab coats or uniforms are seen as more authoritative.
- **Consistency:** People usually feel obliged to uphold a commitment, especially commitments that are made widely known.

In 2016, Cialdini (2016) added a seventh principle of **unity**, which refers to the power of appealing to something the persuader has in common with the audience or a shared

identity. Unity can be used by referring to links in family, education, age, gender, location, religion, or any other common ground the persuader may have with the target audience.

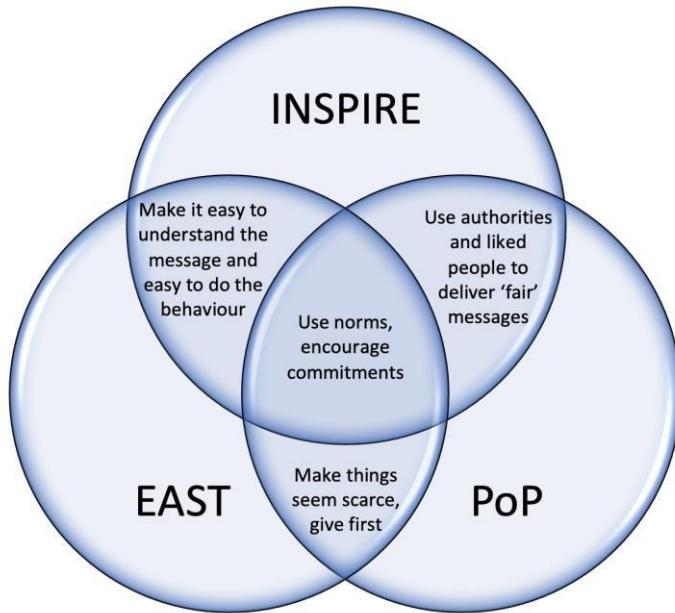
Nearly 10 years ago, BWA applied several of Cialdini's principles in our work on the New South Wales' Home Power Savings Program (Curtis et al., 2017). This program sought to encourage low income households to be more energy efficient. The centrepiece of the program was a visit from a home assessor who spent time looking at bills, appliances and the layout of the home to identify ways in which each household could best save energy. The program utilised:

- **Reciprocity:** Participants were given free items like efficient light globes, power boards and door snakes.
- **Liking:** The home assessors were seen as helpful and trustworthy. Indeed, many people said that they valued their interaction with the assessor more than the free items.
- **Social proof:** BWA trained home assessors to emphasise how many others just like them were choosing the same behaviours.
- **Consistency:** From a shortlist of options, householders were asked to choose up to three behaviours they felt they could do. Householders were then encouraged to write these behaviours down and display them prominently.

WHAT DO MOST GENERIC TOOLS HAVE IN COMMON?

Despite some differences, most generic tools overlap extensively and astute readers will see alignment between many of the tools previously explained (see Figure 3). Consider EAST, where making the behaviour easy aligns with ease from INSPIRE, social (aligns with norms from INSPIRE and Social Proof from PoP) and attractive (overlaps with incentives from INSPIRE). Likewise, in Cialdini's PoP, appealing to scarcity makes a behaviour more attractive (from EAST), social proof is equivalent to norms (from INSPIRE) and making the behaviour social (from EAST), liking and authority aligns with reputation and credibility (from INSPIRE), and consistency has significant overlap with implementation intentions (from INSPIRE) but also makes the behaviour more attractive (from EAST).

Figure 3: Common techniques among generic tools (with indicative examples)



Box 2: Using tailored and generic tools together

Deep-Dive-informed (tailored) interventions and generic behaviour change tools can be used together. Normally we'd suggest first designing a tailored intervention and then adding generic behaviour change tools. Consider, for example, the behaviour of vaccinating against

COVID-19 and that the Deep Dive revealed that it's important to convince people of the effectiveness of the vaccine in preventing hospitalisation. The key message could be delivered by a trusted source (authority) or someone with experience who is similar to the target audience (liking). We could also emphasise that hundreds of millions of people around the world have been vaccinated (norms, social) and make the booking process as easy as possible (ease). In this sense, using several behaviour change tools at once should only increase the chance of success. As researchers, we may not always condone this because evaluating which specific tools worked can be tricky. But as practitioners, the use of these tools in addition to deep-dive-designed tools can be invaluable in achieving desirable behaviour change outcomes.

CONCLUSION

This chapter started with an explanation as to why generic behaviour change tools might be useful as an alternative approach to undertaking a Deep Dive. We typically use these tools in situations where:

1. There are time and budget pressures to deliver an intervention preventing the capacity to undertake Deep Dive activities;
2. Where the reasons for understanding why a behaviour change did or didn't occur are less important.
3. The target audience is large, meaning the method of communication will be less targeted (e.g., convincing all Australians to book a COVID-19 vaccination);
4. The behaviour is not a key focus for you or your organisation; or
5. The behaviour is relatively simple (i.e., a momentary one-off and relatively inconsequential decision).

We outlined three different toolkits that are common in our behaviour change arsenal and showed that there's a lot of alignment between them. The toolkits represent a quick reference guide for behaviour change practitioners to use when one or more of the above situations is relevant. As is often the case, we can only give a brief overview of these approaches and we recommend using some of the additional resources provided to explore further - <https://www.behaviourworksaustralia.org/the-method-book/chapter-9-using-genericbehaviour-changetools/>.

We can't end the chapter without revealing what happened in our example of VicRoads and potentially unfit drivers. We tested whether our new letter worked in a randomised controlled trial, where half of the people receiving a letter got the original letter and half got the new "INSPIREd" one. The simple changes we made to the new letters led to a 12-percentage point increase in on-time submissions (Faulkner et al, 2018). The INSPIRE Framework was then applied to all of VicRoads' medical review letters and, if the results replicate (and we expect they will), on-time submissions will be boosted by 7,000 a year, meaning follow-up correspondence costs will be reduced, fewer people will be unnecessarily suspended and, perhaps most importantly, our roads will be safer to drive on.

In the preceding Deep Dive chapters, we've showcased influences on behaviour, methods for understanding them, and matching audience insights to interventions. To this, we've added some generic tools which can be either used separately (in the absence of a Deep Dive) or in addition to interventions that emerge from Deep Dive processes. At this point, those following The Method should have one or more evidence-based behaviour change tools that they think might work to change behaviour. These tools may have emerged from literature and practice reviews based on 'what works', from the Deep Dive approaches articulated previously, or from this chapter. But either way, you are ready to implement an intervention. This leads us into the next phase of The Method - *Application*. The Application phase outlines tools that can be used to determine if your intervention worked, with the next Chapter focusing on how to identify potential measures of success.

REFERENCES

- Brenan, M. (2017). Nurses keep healthy lead as most honest, ethical profession. *Gallup*, 2021(26 August).
Retrieved from <https://news.gallup.com/poll/224639/nurses-keep-healthy-lead-honestethical-profession.aspx>
- Castelo, N., Hardy, E., House, J., Mazar, N., Tsai, C., & Zhao, M. (2015). Moving citizens online: Using salience and message framing to motivate behavior change. *Behavioral Science & Policy*, 1(2), 57-68.
- Cialdini, R. B. (2009). *Influence: Science and practice*. 5th Ed. Boston: Pearson Education.
- Cialdini, R. (2016). *Pre-suasion: A revolutionary way to influence and persuade*. London: Cornerstone.

Curtis, J., Graham, A., Ghafoori, E., Pyke, S., Kaufman, S., & Boulet, M. (2017). Facilitating adaptive management in a government program: A household energy efficiency case study. *Journal of Environmental Management*, 187, 89-95.

Dal Bó, E., Finan, F., & Rossi, M. A. (2013). Strengthening state capabilities: The role of financial incentives in the call to public service. *The Quarterly Journal of Economics*, 128(3), 1169-1218.

David, M., & Ware, R. (2014). Meta-analysis of randomized controlled trials supports the use of incentives for inducing response to electronic health surveys. *Journal of Clinical Epidemiology*, 67(11), 1210-1221.

Faulkner, N., Borg, K., Bragge, P., Curtis, J., Ghafoori, E., Goodwin, D., Jorgensen, B.S., Jungbluth, L., Kneebone, S., Smith, L., Wright, B. and Wright, P. (2019). The INSPIRE Framework: How public administrators can increase compliance with written requests using behavioral techniques. *Public Administration Review*, 79(1), 125-135.

Faulkner, N., Jorgensen, B. S., Sampson, J., & Ghafoori, E. (2018). Improving compliance with medical fitness to drive reviews: The role of behaviourally-optimised letters. *Transportation Research Part F: Traffic Psychology and Behaviour*, 55, 1-8.

Gans, K.M., Risica, P.M., Strolla, L.O., Fournier, L., Kirtania, U., Upegui, D., Zhao, J., George, T. & Acharyya, S. (2009). Effectiveness of different methods for delivering tailored nutrition education to low income, ethnically diverse adults. *International Journal of Behavioral Nutrition and Physical Activity*, 6(1), 1-11.

Goldstein, N., Martin, S., & Cialdini, R. (2010). *Yes!: 50 Scientifically Proven Ways to Be Persuasive*. New York: Free Press.

Mantzari, E., Vogt, F., Shemilt, I., Wei, Y., Higgins, J., & Marteau, T. (2015). Personal financial incentives for changing habitual health-related behaviors: A systematic review and meta-analysis. *Preventative Medicine*, 75, 75-85.

Milkman, K., Beshears, J., Choi, J., Laibson, D., & Madrian, B. (2013). Planning prompts as a means of increasing preventive screening rates. *Preventative Medicine*, 56(1), 92-93.

Milkman, K., Beshears, J., Choi, J., Laibson, D., & Madrian B. (2011). Using implementation intentions prompts to enhance influenza vaccination rates. *Proceedings of the National Academy of Sciences*, 108(26), 10415-10420.

Milne, S., Orbell, S., & Sheeran, P. (2002). Combining motivational and volitional interventions to promote exercise participation: Protection motivation theory and implementation intentions. *British Journal of Health Psychology*, 7, 163-184.

Phua, J. (2016). The effects of similarity, parasocial identification, and source credibility in obesity public service announcements on diet and exercise self-efficacy. *Journal of Health Psychology*, 21(5), 699-708.

Schultz, P., Nolan, J., Cialdini, R., Goldstein, N., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, 18(5), 429-434.

Scott, P., & Edwards, P. (2006). Personally addressed hand-signed letters increase questionnaire response: A meta-analysis of randomised controlled trials. *BMC Health Services Research*, 6(1), 111.

Service, O., Hallsworth, M., Halpern, D., Algare, F., Gallagher, R., Nguyen, S., . . . Kirkman, E. (2014). *EAST: Four Simple Ways to Apply Behavioural Insights*. London, UK: The BIT

Wenzel, M. (2006). A letter from the tax office: Compliance effects of informational and interpersonal justice. *Social Justice Research*, 19(3), 345-364.

APPENDIX 1: THE INSPIRE FRAMEWORK

Technique	Description	How to apply
I mplementation intentions	Implementation intentions use 'if-then' plans <ul style="list-style-type: none"> - that link situational cues with goal-directed - action responses. 	Ask readers to write down 'when and where' (time of day/day) they intend to undertake actions Insert boxes in part of the letter where recipients can list the day, date, and time they will perform the relevant behaviour
N orms	Injunctive norms refer to beliefs about what is socially approved or disapproved of by others. Descriptive norms refer to beliefs about how most other people actually behave.	<ul style="list-style-type: none"> - Provide feedback on behaviour of letter recipient compared with peers/or others in similar situations, especially when the comparison is with good behaviour performance. e.g., 9/10 people return their tax on time Carefully select a comparison group e.g., energy efficient neighbours - Avoid the boomerang effect by including a positive reinforcement (e.g., smiley face) when recipient is already performing the desired behaviour well
S alience	<ul style="list-style-type: none"> - Properties that make a letter stand out and attract attention, and differentiate it from other similar letters. 	<ul style="list-style-type: none"> - Use different coloured backgrounds - Provide handwritten signatures - Use recipients name - Use registered post -- Stamp with red "pay now" notice - Use multiple stamps
P rocedural justice	<ul style="list-style-type: none"> - Procedural justice emphasises the fairness of a process, including: provision of accurate information about the process; lack of bias; - 	<ul style="list-style-type: none"> - consistent procedures; and respectful treatment. - Use communications that are polite, respectful, and written in a manner that preserves the dignity of the recipient Provide reasoning of why the recipient has received the letter - Explain the procedures as they apply to everyone - Explain recourse options

Incentives

Receiving a beneficial consequence (reward) -
in response to a particular behaviour increases -
the likelihood of that behaviour. Incentives -
motivate individuals to perform a specific
behaviour.

-

-

Reputation

The reputation or credibility of the source of
the information is an influential factor in
communications and can be used to further
persuade readers.

-

-

Ease

Make compliance with instructions easy.
Ensure the letter is easy to read, the written
instructions are clear, and make the target
behaviour straightforward.

-

-

-

-

Offering monetary incentives can be effective

Offering non-monetary (donations, prize draws, gifts) incentives is more effective than no incentive

Match incentive size to size of request

Inform readers of the negative consequences of inaction (disincentives) Check that incentives don't 'crowd out' intrinsic/social motivations

Appropriate for one-off behaviours rather than sustained change

Consider who the message should come from

Present source before the message Include credibility from an in-group member

Include supportive quotes from credible experts

Make use of content hierarchy; place most important information at the top and relegate details to an information sheet Check readability of text

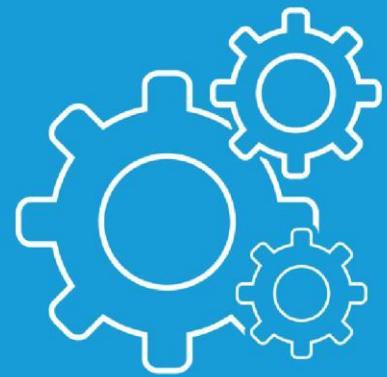
Use sufficient text size

Include headings and subheadings

Include positive instructions in an active voice

Avoid acronyms and jargon

- Judicious use of diagrams
 - Divide complex requests into small steps
-



THE METHOD

CHAPTER 10: How to choose good success measures

By Stefan Kaufman and Bernice Plant

www.behaviourworksaustralia.org



INTRODUCTION

Wendy and her colleagues work for a regional organisation of local governments. They were grappling with how to measure success in waste avoidance programs across councils, and asked BehaviourWorks Australia (BWA) to provide advice and suggestions on how they could better measure both success in behaviour change (a key intermediary outcome), and in waste (their impact area). They had reason to believe that good examples were scarce in waste avoidance, and wondered if solutions existed in other policy areas. While there was no simple fix, what we learned from working with them underlines the importance of wellconsidered and measured success measures in behavioural public policy. Success measures can mean very different things in different contexts. In nearly every case though, it involves lifting our eyes from tracking progress, implementation and activity, to whether or not we are delivering the intended changes that motivate and justify an initiative, AND in what way we can reasonably conclude our efforts contributed to those changes.

The previous chapters of the BWA Method Book have taken you through the Exploration and the Deep Dive phases, which have provided some useful tools to understand your problem, arrive at a target behaviour of interest, identify key behavioural influences, and design potential interventions. But how do you know whether your intervention is effective (or, said another way, ‘*What works?*’) – how do you measure and assess the impact of an intervention and prepare to scale it up? These questions are addressed during the Application phase of the BWA Method. The remainder of the book therefore aims to provide you with some useful tools to help you with assessing the impact of your intervention and putting your intervention into practice. This chapter addresses a key starting point: how to identify appropriate measures of success.

HOW ARE SUCCESS MEASURES ARTICULATED?

Defining measures of success is typically the first step to assessing impact. Several related and important terms, concepts and tools are needed to clearly articulate success ‘indicators’, and track them well as ‘measures’:

Evaluation: An evidence-based judgement of the merit or worth of a program or activity.

Theory of change: Most simply, this is our articulation of ‘how and why our initiative makes a difference’ (Weiss, 1999). ‘How’ can take the form of hierarchies of differences or changes – a very generic ‘spine’ for a theory of change might specify: inputs; activities; outputs; participants; short-, medium-, and/or long-term outcomes; and impact. They can take the form of a relatively simple logic, a complex program hierarchy, or other tools that link program activity to different outcomes and impacts. The ‘why’ is any specification from theory or research of mechanisms that are the core ‘special sauce’ of an initiative, and can be problem and sector specific (e.g., ‘restorative justice’ might apply to a community reintegrative justice program), while others are relatively generic (e.g., Ajzen’s 1991 ‘theory of planned behaviour’ is widely applied across a broad range of contexts and behaviours - see Funnell & Rogers, 2011).

Situation assessments: A helpful early step in articulating theories of change and appropriate indicators and measures is a situation assessment. This can usefully include the ‘spine’ of a very generic theory of change in a table with questions asked at each level of the spine about ‘what are our starting conditions’. It can also draw on a system map (see Chapter 2). Laying out observations about the pre-conditions of an intervention in this way can help articulate the

'failure' or 'problem' indicators or measures you need to flip into success measures. Or to put it another way: 'There is no such thing as a dysfunctional organisation, because every organisation is perfectly aligned to achieve the results it gets' (Heifetz, Grashow, & Linsky, 2009).

Indicators: Are often narrative or stated changes embedded in the theory of change that are known (ideally) or believed (more often) to be crucial for the success of a program or initiative. They often attempt to articulate the true nature of a change that is desired and are commonly articulated as a change for a person or thing, in a time and/or context.

Measures: Are specific data and analyses measuring an indicator, and can be multiple and varied, addressing the same indicator. In rare cases, the indicator and measure can be stated as the same thing. More often, an individual measure approximates – or multiple measures 'triangulate' – the indicator.

In Appendix 1, we present a tool that was first introduced to BWA by Professor Stephen Montague, a visiting researcher from Canada, as a resource to support articulating success measures in government programs. We like it because if you've done evidence reviews (see Chapter 1), systems mapping (see Chapter 2), Data (chapter 3) and stakeholder consultation (see Chapter 4) and behaviour prioritisation (see Chapter 5) in the Exploration phase of the BWA method, you are in a great position to articulate a **situation assessment** and **ultimate success indicators and measures** that motivate the intent of your program.

Similarly, most of the steps of the Deep Dive phase (see Chapters 6-9) help a great deal with articulating **intermediary indicators and measures** that, if changed, are known (ideally) or are believed to be likely to impact ultimate success indicators and measures via your **theory of change**. Specific datasets and ideas for measures for the ultimate impact might emerge from an understanding of the 'problem', while key success measures can be usefully-tied to the drivers and barriers to behaviour change that you identify, or tied to what you prioritise behaviours against. Indeed, foundational audience research, such as adoption levels of a given behaviour, may also be a baseline measure for some indicators.

Finally, assessing proposed interventions against the intended impact can help distinguish the differences between what works to change behaviour, versus what works to address the reason you want to change it. This can help avoid trialling or scaling trivial initiatives or those with unintended disbenefits. For example, reusable shopping bags as an alternative to single use plastics only break even on a lifecycle impact basis if used many times, depending on the type. If you identify that people are more likely to drive around with ten cotton bags – having bought new ones because they are embarrassed each time they forget – you might reconsider how you roll out the intervention, or whether or not you should encourage cotton at all (hot tip: you probably shouldn't). Similarly, if you are more concerned about littering, greenhouse gas emissions or resource consumption, you might prioritise different types of bags or something else entirely¹⁰. For these reasons, despite success measures appearing in the Application phase of The Method, you actually need to be thinking about and working on your success measures right from the outset. They are almost impossible to articulate and to measure *well* in the end stages of a program – as one of Australia's leading evaluators, Professor John Owen, observed: '80% of a successful evaluation is a well-designed program' (Owen, 2016).

¹⁰ A useful summary of shopping bag lifecycle assessments can be found here
<https://ourenvironment.ac.nz/2018/09/10/howsustainable-is-your-grocery-bag/>

Returning to our opening case study

An early task with Wendy and her colleagues was reviewing over 60 waste avoidance programs delivered across their council areas. One part of this was assessing with the steering group where some of the ‘pain points’ in evaluating waste avoidance outcomes were. Figure 1 outlines this discussion with observations of pain points in the bolded comment boxes. Further interviews with council staff identified that many programs were being run either for good reasons, but not prioritising behaviour change and waste avoidance outcomes, or were being run out of organisational routine and habit. Some examples were supported by external frameworks and tools from funders and collaborators, and had more developed success measures, but sometimes struggled to link these back to relevant decision making in the program itself. Our advice in the end was essentially ‘there are many

good reasons to run programs that don’t aim to change behaviour or impact waste avoidance, but decide which programs for which these two aspects are the defining value, and then, if necessary, (re)design them to do it’.

We advised this knowing that measuring and evaluating success is a lot more worthwhile and achievable when, as a former CEO of one of the authors liked to say: ‘We know where we are going’, ‘We know how we are going to get there’, and ‘We can tell how far we’ve come’. Of course, we also acknowledged then, and here, that this can be easy to say and hard to do. Our interviews also highlighted limited capacity and expertise for behaviour change, design and evaluation, inconsistent funding and staffing, and other fundamental challenges. This highlights the importance of considering implementation and scaling, capability building and other aspects addressed in the following chapters.

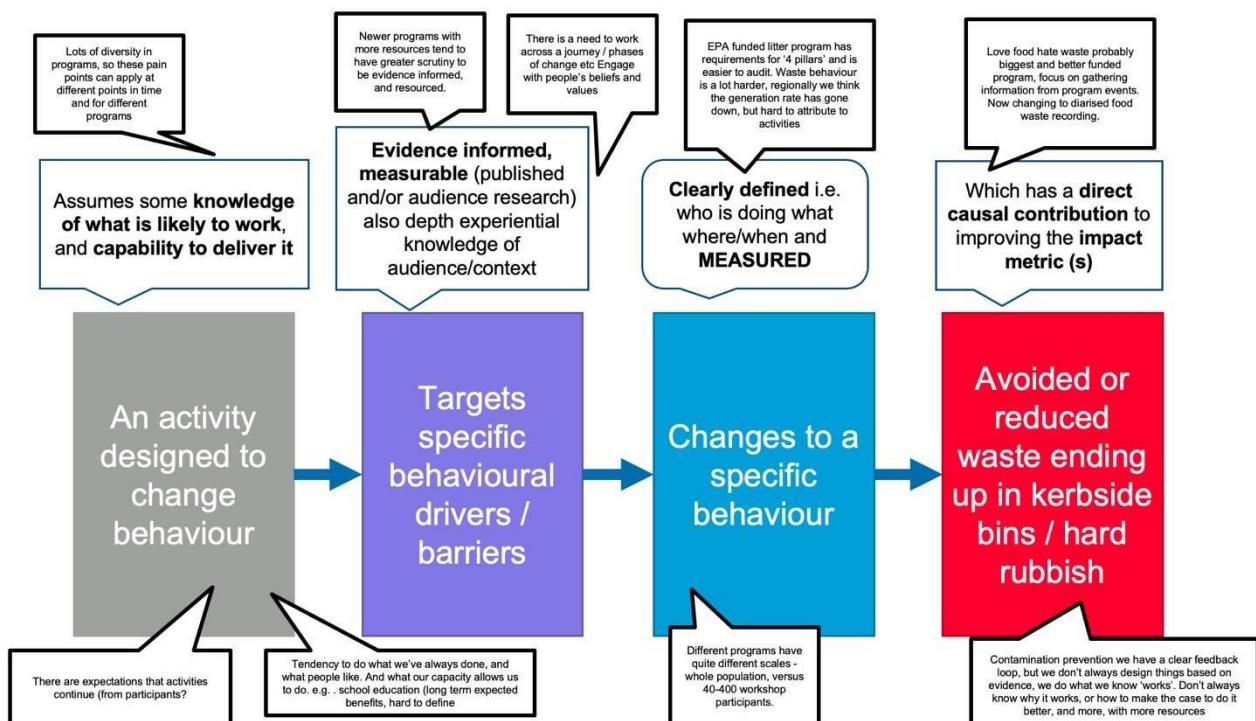


Figure 1: Pain points in evaluating waste avoidance outcomes

A key consideration in articulating success measures

When considering and articulating success measures, it is important to identify and consider the needs of different program stakeholders in defining success, including during implementation. Evaluator Jess Dart sometimes refers to the tension between ‘Evaluation for us, versus evaluation for them’ – highlighting that the needs and interests of program deliverers, their bosses, program beneficiaries/end users, and program funders can all be different. For instance: i) the *program team* might want a safe space to implement the program, generate feedback, innovate, or reflect on nutting out a difficult challenge; ii) *their bosses* might want popular programs that run smoothly; iii) the *end users* might want effective tools that make a positive difference to experienced problems, and quickly; and iv) the *funders* might want accountability and fidelity to the originally funded activity, towards a grand but distant goal.

As you can imagine, the different needs of the various stakeholders involved can create tension. Unintegrated success measures that are easy to measure, but trivial (e.g., the number of flyers we distributed) – versus those that are important, but hard to move, measure, and long term (e.g., the reduction of tonnage of waste per capita) – can exacerbate conflicting understandings of success, and can motivate misplaced effort, the ‘gaming’ of indicators, and worse. Indeed, ‘Goodhart’s law’ rather pessimistically warns that as soon as a measure becomes the basis of success, and particularly the basis of incentives and rewards, it becomes suspect (Manheim, 2016).

But it doesn’t have to be this way. Recognition of such tensions has also led some funders and program delivery organisations to reimagine and redesign the role of success measures, evaluation, and reporting. See for example reflections from the World Bank Behavioural Insights team (World Bank, 2015), and Professor Kim Jim, in the Conclusion.

Articulating your own success measures

Considering all the above, and drawing on earlier phases of the BWA Method, you may find you can answer many of the questions in the table in Appendix 1 without further research and consultation, although the completed table can be a useful resource for communicating the rationale and evidence base of the program. Where you can’t answer the questions, this will also help you to identify gaps in your knowledge that could be prioritised. Of course, not all programs and theories of change fit the structure shown in Appendix 1 (based on ‘Bennett’s Hierarchy’, which assumes changing hearts, minds, and behaviour are central, see Rockwell & Bennett, 2004), so feel free to modify this to fit other logics that make more sense for your initiative. Different theory of change ‘why’ mechanisms may also suggest specifying different elements in the rows, and different row orders. For example, where target behaviours are considered habitual, knowledge and skills change may be less important to track than implementation of efforts to restructure environments.

Any given initiative will almost never address each and every question – and certainly not with comprehensive measures. Rigorous measurement is often difficult, and expensive, and like any other aspect of the intervention, tradeoffs will need to be made. Considering these questions should help you identify important indicators. Looking at the following list, more often than not, the further down this list you go (see also Figure 2), the harder to measure they are, but the more important they are to measure. So, it’s important to choose them carefully, and not to stop at the easiest measures:

Capability, inputs and outputs (i.e., how much of what kind of effort are you producing). These can be complemented by immediate **Reach** indicators – i.e., what proportion of your target audience are you reaching, who is participating etc. – and **Reactions** indicators – for example of user experience, event satisfaction etc. Generally speaking, well run organisations have most of these covered. It’s the next layers where you will increasingly need expert capability (whether in-house or with researchers).

Antecedents of behaviour change (i.e., known drivers of change, that if achieved, are likely to lead to behaviour adoption). In technical terms, ‘behavioural mechanisms and mediators’ – e.g., having attended the event, did participants increase their desire to act or in their confidence in their ability to do so?

Lead indicators of the adoption of the behaviour change itself, often measurable during or soon after the intervention (but not always), and preferably compared with a baseline or control group. Defining and measuring behaviour is not simple - (see *TIP* below), and previous chapters.

Lag indicators of the desired impact motivating the initiative, often measurable later, and with more confounding and contributing factors. These are the ‘ambient conditions’ and are often closely aligned with the mission or purpose of your organisation or initiative.

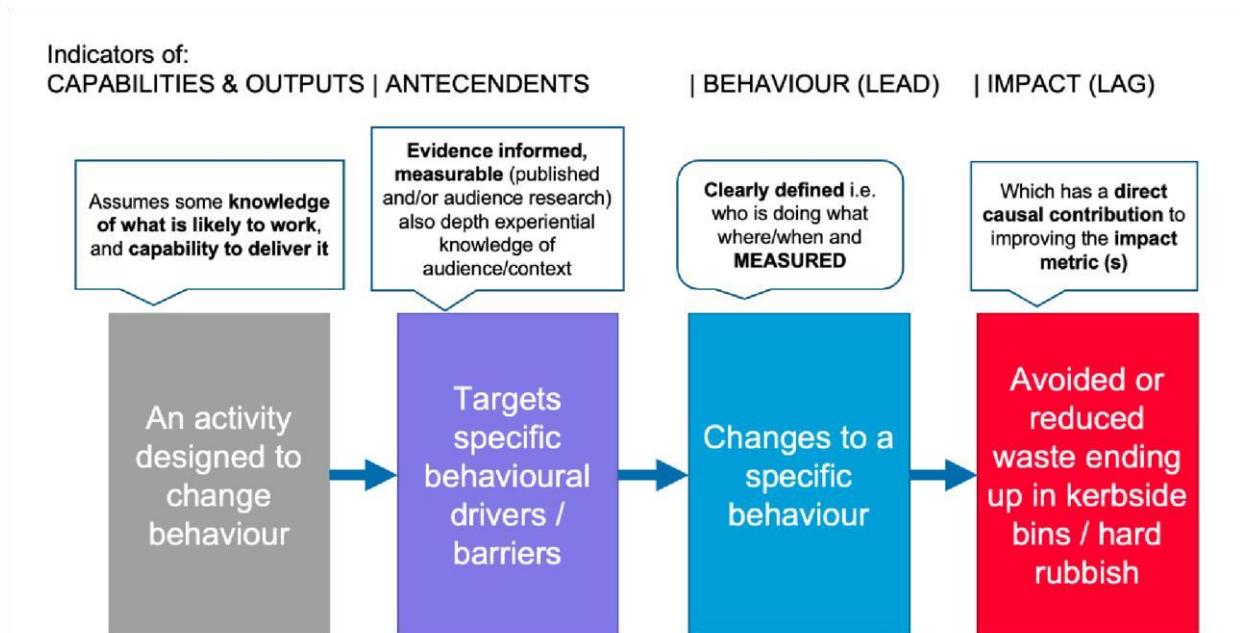


Figure 2: Examples of different indicators to measure the impact of a waste avoidance intervention

TIP: Typically, behavioural scientists have a hierarchy of evidence when it comes to measuring behaviour:

- **Direct observation** - overt or covert (e.g., did people litter or not)
 - **Inferred by indirect measures** (e.g., energy consumption meter data matched to appliance use)
 - **Self-reported current behaviour** (where well crafted, salient questions can support recall; hard to do well in a survey)
 - **Self-reported intended behaviour** (supported by well-crafted questions and relevant, specific scenario; hard to do well in a survey)
-

CASE STUDIES: EXAMPLES OF HOW WE HAVE CONSIDERED AND SELECTED MEASURES OF SUCCESS

In this section, we present two case studies to illustrate some of our processes and considerations when selecting measures of success. Here, we draw upon two trials that were completed as part of the Patient Safety Research and Innovation Program

(behaviourworksaustralia.org/victorianmanaged-insurance-authority) in partnership with the Victorian Managed Insurance Authority (VMIA). The different approaches of these trials provide an opportunity to highlight some key considerations that could be useful for others when selecting their own measures. To this end, each case study includes an overview of: the trial background and objective; the key considerations in selecting measures and the process by which we reached them; some examples of key measures we arrived at and a few key lessons that we took away from the trial, which could help to guide the selection of appropriate measures for future trials that aim to assess effectiveness and/or the feasibility of interventions. In describing sample measures used for each case study, we attempt to align these with the key Theory of Change components in the table presented in Appendix 1.

VMIA – Escalation of Care Trial

Trial background and objective

Early escalation of patient deterioration can help improve patient outcomes. However, there are a range of behavioural barriers to speaking up when something is ‘not right’ with a patient. An immersive ‘choose-your-own-adventure’ training video that aimed to help clinicians gain experience in navigating some of the key challenges around having escalation of care (EOC) conversations was trialled at two Victorian hospitals. The video, developed by The Shannon Company, was designed to help clinicians gain experience in navigating some of the key behavioural barriers to escalation of care conversations not breaching the Medical Emergency Team (MET) criteria (i.e., during the pre-MET period).

We’ve chosen to draw on this case study because of the diverse range of measures and methods used, including the use of daily experience sampling, in the context of a trial conducted in two Victorian hospitals.

Key considerations in selecting measures of effectiveness and feasibility

As mentioned earlier in this chapter, a key consideration when articulating success measures is the needs of different program stakeholders in defining success. As such, at BWA, we typically arrive at measures of success in consultation or close collaboration with key stakeholders (indeed, you may have noticed this theme in previous chapters, as there is great value in stakeholder engagement across all stages of the BWA method!).

For the EOC trial, we arrived at our success measures in partnership with key stakeholders, including VMIA, SaferCare Victoria, clinicians, clinical educators, and researchers. Some of our main considerations in identifying and selecting (or developing our own) success measures for this trial included the: alignment with the trial and intervention objectives; feasibility within the constraints of the trial; and that there was a goal to submit the findings from the trial for publication. For example: Although, from a behaviour-change perspective, we would consider ‘observable’ behaviour (e.g., observing actual EOC conversations) higher in the hierarchy of measures (see Appendix 1), an important consideration in selecting our measures was that it wasn’t *feasible* – for ethical and practical reasons – for us to observe EOC conversations in hospitals. This meant that

when it came to measuring our lead indicators, we arrived at self-report measures. Another example of such a consideration for this trial is: Given that an objective was to submit the findings for publication, it was a priority to draw upon pre-existing measures with demonstrated validity and reliability, where possible (see Ruane, 2004 for thorough discussions on the aspects of validity and reliability of measurement). Although the majority of measures employed during this trial were designed specifically for this trial (see *TIP* below).

TIP: *Given the diverse nature of behaviour-change projects, finding valid and reliable measures that align with specific contexts for a given project can be challenging, and so adapting measures used in other settings to fit the context of your project and/or drawing on other evidence to develop your measures may be required. Ultimately, balanced consideration is needed here: it may be better to design a ‘new’ measure specifically for your project that is a better ‘fit’ with your project objective than to use an existing, validated measure that is poorly aligned with your objective. Where ‘new’ measures are specifically designed, we would strongly recommend taking an evidencebased approach to developing items that tap into the constructs of interest (e.g., using insights gained via evidence reviews and/or stakeholder consultation, see Chapters 1 and 4, respectively).*

What we measured and key lessons we took away for future trials

The aim of the trial was to pilot test the *effectiveness* and the *feasibility* of the interactive ‘chooseyour-own-adventure’ training video as a potential intervention to address some of the key barriers involved in escalation of care conversations during the pre-MET period. This meant that a range of measures were required to assess effectiveness and feasibility, including:

- Lead indicators / Intermediate outcomes / target audience behaviour (e.g., reported willingness/reluctance to escalate patient care and reported number of EOC conversations to measure lead indicators);
- Antecedent indicators / immediate outcomes / behavioural antecedents (e.g., reported confidence to escalate patient care, and knowledge of how to escalate care to measure antecedent lead indicators); and
- Reactions and inputs/capability (i.e., reported advantages, disadvantages, and suggested changes to the intervention to measure experiential and service delivery indicators).

A range of methods were applied to gather this data, including two online surveys (at the start and end of the trial period) and daily experience sampling, gathered via participants’ mobile phones (throughout the trial period). The use of experience sampling could be considered a strength of the data collection methods used in this trial – if gathered at regular intervals (e.g., daily, as was the case for this trial), responses gathered via experience sampling can be less susceptible to failures in memory or recall biases than self-reported data gathered over longer time periods. **Inter-hospital transfers trial**

Trial background and objective

Inter-hospital transfers (IHT) play an important role in the healthcare system, ensuring patients can access appropriate care in a timely manner. However, difficulties in locating available beds and pushback around the necessity of transfer are some of the biggest identified barriers to enabling

transfers. A persuasive video was trialled, which aimed to improve collaboration between sending and receiving doctors during non-critical IHT request phone calls. The video, developed by The Shannon Company, was designed to foster an empathetic and collaborative mindset among doctors who receive non-critical patient transfer requests from a sending hospital.

We've chosen to draw on this case study because of the indicators of collaborative behaviour that were measured within the constraints of an online trial and because this trial included measures selected to provide insight into potential mechanisms of change by the intervention.

Key considerations in selecting measures of effectiveness and feasibility

As for the previous case study, we arrived at our selected measures for this trial in partnership with key stakeholders, including VMIA, SaferCare Victoria, clinicians, and researchers. Our main considerations in identifying and selecting (or developing our own) measures of success for this trial were similar to that described for the EOC trial (above), which included the: alignment with the trial and intervention objectives; feasibility within the constraints of the trial; and that there was a goal to submit the findings from the trial for publication. However, while these priorities were similar to the EOC trial described above, their implications were quite different within the context of the IHT trial. For example, given that the trial needed to be conducted online – again, recording or observing phone conversations was not feasible for practical reasons nor from an ethical perspective. This raised a challenge for us in attempting to observe changes in ‘collaborative behaviour’ – we attempted to infer collaborative behaviour by measuring the number of questions and additional steps that participants offered the ‘caller’ during the trial (i.e., we used audio vignettes to simulate non-critical IHT request phone calls, and participants could type questions and steps into the online survey).

What we measured and key lessons we took away for future trials

The aim was to pilot test the *effectiveness* and *feasibility* (in the form of its perceived acceptability to members of the target audience) of a persuasive video as a potential intervention to improve collaboration between sending and receiving doctors during non-critical IHT request phone calls. A range of measures were required to assess effectiveness and acceptability within the constraints of an online trial, including:

- Lead indicator / Intermediate outcomes / target audience behaviour (e.g., reported likelihood of accepting the patient and the reported number of questions and additional steps that would be offered to the caller to measure lead indicators);
- Antecedent indicator / immediate outcomes / behavioural antecedents (e.g., reported perceptions of the scenario, including the perceived severity and urgency of the medical case to measure antecedent lead indicators; reported affective and cognitive empathy towards the caller to measure potential behavioural mechanisms and mediators); and
- Reactions and inputs/capability (e.g., perceived impact of the video, and what they liked and what they would change about the intervention to measure experiential and service delivery indicators).

Many of the broader lessons we learned by conducting this trial apply to trial design (covered in a future chapter); however, some of our reflections on the measures used in this trial are worth noting for future trials. Although not equivalent to dynamic conversations that may occur over the phone, the use of open-ended question formats for participants to list additional questions and steps that they would take with a caller was an innovative approach that allowed us to infer collaborative behaviour

that could occur during a phone conversation. Some of the measures used in this study were adapted from pre-existing measures, which strengthened the validity of measurement for the constructs of interest.

TIPS AND ADVICE FOR ADVANCED APPLICATION

Reflecting on the case studies provided in the previous section, it is perhaps worth mentioning some trends in this work. There tends to be (relatively speaking) a heavier focus on the combination of intermediate outcomes (target audience behaviour), immediate outcomes (behavioural antecedents), and reactions. Ultimate or end outcomes may be measured where possible, but they are often not feasible or within scope for our projects for various reasons (i.e., as they typically require a longer duration). Instead, we may select target behaviour(s) for which there is a known relationship with the outcome of interest – we may make inferences about the potential impact of the intervention on outcomes via their measured impact on behaviour.

A broader reflection on success measures used in behaviour-change projects is that while measures are typically selected that help to address the question “*What works?*”, relatively speaking, it is less common for measures of potential behavioural mechanisms or mediators to be included. This then limits our understanding of *how* particular interventions work, which is a limitation not only for the project of interest, but for future refinement and development of interventions and for contributions to theory within the field (also see discussions about process evaluations in Abraham & Denford, 2020). One of the big challenges, however, in measuring behavioural mechanisms is how to include such measures without compromising the fidelity of the trial and/or the measurement of lead indicators – particularly when it comes to conducting field trials. Again, this is where balanced consideration is needed regarding the project’s priorities, also noting that it may be possible to gain insights into *how* something works via the design of your trial and your intervention conditions (i.e., systematically varying aspects of an intervention vs. testing ‘cocktail’ interventions) or via subsequent experimental research. Some of these ideas will be discussed in the chapter on trial design.

Once again, you will never be able to measure everything, but thinking through the complete suite of possible measures and making careful choices is time well spent. We will leave you with a few guiding questions and their implications for selecting measures of success:

Primary considerations when selecting measures:

- How well do the measures align with the goals of the intervention and the trial? (The research question)
- How well do the measures align with the outcomes of interest? (The problem)
- How feasible are the measures, considering any constraints around trial design and resources available (e.g., outcomes often require longer project durations; interviews are labour-intensive)?
- Are there any additional priorities that need to be considered regarding the measures, from the perspective of: partner organisation, target audience, society?

Additional considerations when selecting measures of success:

- Is there a goal to publish (i.e., validity, reliability, bias of measurement)?
- Is there any existing data available that you can tap into (e.g., from partner organisations)?
- What measures might allow you to compare with existing data / standards in the field or in other contexts?

- What measures are needed to gauge the representativeness of your sample?
- What additional measures might be needed to complete the picture or add value (e.g., controlling variables, moderators, mediators)?
- What disciplines, theories and methods have been used, at each level, and are they relatable and coherent (i.e., inter-disciplinary integration)?

Important considerations from an ethical perspective when selecting measures:

- Participation (have you involved everyone with a legitimate interest, and especially those who lack the power or resources to involve themselves, in defining and measuring success? Are the measures meaningful and relevant to them?)
- Potential identifiability (e.g., surprisingly few combinations of demographic information is needed to be able to potentially identify someone)
- Consent (e.g., covert observation raises ethical risks; publicly available information, does not necessarily mean that this information is available for use in your project)
- Burden (e.g., time to complete your measures – balancing completeness of data with the burden on participants)
- Potential psychological harm (e.g., considering the nature of topics included in your measures, particularly if you are working with vulnerable participant groups)
- Legal and reputational risks (e.g., to participants, but also from the perspective of your organisation and/or client in reporting the findings)

CONCLUSION

Our advice to the regional local government organisation we introduced at the beginning of the chapter certainly wasn't intended to say they were doing the 'wrong things'. Rather it was if they really want to prioritise behaviour change that leads to avoided waste, programs need to be designed and implemented with a focus on doing that. There are many other good reasons to run programs with a less direct focus – e.g. to create a supportive environment for change by raising awareness, informing and educating, and building skills; to provide a service participants value and enjoy; to build networks and social inclusion, and none of these goals are incompatible with a behaviour change program (indeed they may be crucial for them). At the time of writing, we are exploring developing a training program to help them redesign initiatives to target waste avoidance behaviour change and reduced waste, where this is a priority, so watch this space. A key focus will be tightening the threads between the 'why' of initiatives, intended behaviour changes, and how these are supported – the line of sight between activity, behaviour and impact. We provided a generic theory of change similar to the tool introduced here to help with this.

This example is symptomatic of a broader challenge. 'Pro-social' organisations (e.g., government and NGOs, for-cause businesses etc.) are often accountable to show the efficiency of outputs (i.e., workshops) for inputs (i.e., funding) while pursuing important, big picture goals. Whether or not such accountability actually supports achieving the pro-social good objective is often doubtful. A narrow focus on output efficiency makes good decision making about what to keep doing, start, stop or do differently very hard, not to mention leaving staff wondering if they are treading water instead of kicking goals. It also leaves funders and policy organisations unsure if they are directing efforts in the right direction, and leaves people receiving services, or in need, doubtful if their needs are being met. Indeed, former World Bank Chair Professor Kim Jim explained that part of the reason he was interested in behaviour change approaches was to help his own organisation, and particularly leaders, define success measures, and make better decisions (Dubner, 2019):

Sometimes people say, ‘The private sector does everything better.’ I don’t know that that’s really the case so much as [those] private sector entities that did it poorly no longer exist because they go out of business. Public sector entities can stay in business for a very long time no matter how poor their performance is ... I’ve been trying to understand — in the absence of market forces — how can you raise the temperature so that people really focus on improving execution? Because in the public sector, not only do we tolerate poor execution, but often, unfortunately, we celebrate poor execution. Poor execution, sometimes, for people is a symbol of the fact that you’re public and not private sector (Prof. Kim Jim interviewed in Dubner, 2019).

The Application phase of The Method Book is not just about working out ‘what work’s, but also ‘why’, for ‘whom’ and ‘how much’, and following through with the implications of this for effective public policy and programs. This requires assessing if you are asking the right questions, and learning effectively. But to answer such questions, we need appropriate measures of success (hence it is typically the first step in the Application phase). And the measures need to cover a spectrum of considerations across not just what changes behaviour, but what supports effective, efficient and appropriate interventions, service and program delivery and more. For these reasons, systems change approaches prioritise indicators and measures of learning, capability building, and adaptability over strict adherence to predefined success measures, to better take into account systemic relationships and to drive fundamental change. But it’s very hard to steer towards ultimate success without meaningful feedback and intermediary measures. Having a well-articulated behaviour change rationale for a program, policy or service can help make behavioural aspects of systems change much more measurable, providing a crucial ‘missing middle’ for more adaptive theories of change that hinge on people doing something differently (i.e., our goal is to transform X, and to do that we think we need to see behaviour change Y to happen; if it doesn’t, our goal is still X, and we will learn from the effort and try something else). And let’s face it, how many things that are worth doing don’t involve trying to get someone to act differently?

Now that you are armed with your preferred interventions and measures of success, you are now in the position to design and implement trials to explore the impact of your interventions on your measures of success. That’s the focus of the next chapter.

REFERENCES

- Abraham, C., & Denford, S. (2020). Design, implementation, and evaluation of behavior change interventions: A ten-task guide. In M. S. Hagger, L. D. Cameron, K. Hamilton, N. Hankonen & T. Lintunen (Eds.) *The Handbook of Behavior Change*, 269-284. Cambridge, UK: Cambridge University Press.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2), 179211. doi: 10.1016/0749-5978(91)90020-T.
- Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory: Effective use of theories of change and logic models* (1st ed.). San Francisco, CA: Jossey-Bass.
- Heifetz, R., Grashow, A., & Linsky, M. (2009). *The practice of adaptive leadership: Tools and tactics for changing your organization and the world*. Boston, MA: Harvard Business Press.
- Manheim, D. (June, 2016). *Goodhart’s Law and Why Measurement is Hard*. Ribbonfarm: Constructions in magical thinking Website, accessed via <https://www.ribbonfarm.com/2016/06/09/goodharts-law-and-why-measurementis-hard/>

Owen, J. (2016, September 17–21). *The landscape of evaluation theory: Exploring the contributions of Australasian evaluators* [Conference presentation]. Australasian Evaluation Society Conference 2016, Perth, Australia.

Rockwell, K. & Bennett, C. (2004). Targeting outcomes of programs: A hierarchy for targeting outcomes and evaluating their achievement. *Faculty Publications: Agricultural Leadership, Education & Communication Department*, 48, accessed via <https://digitalcommons.unl.edu/aglecfacpub/48>

Dubner, S. J. (Host). (2019, January 12). Hacking the world bank (Ep. 197 Update) [Audio podcast episode]. In *Freakonomics*. Accessed 30 September 2021 via <https://freakonomics.com/podcast/jimyong-kim-update/>

Ruane, J. M. (2005). *Essentials of research methods: A guide to social science research*. Malden, MA: Blackwell publishing.

Weiss, C. H. (1999). Nothing as practical as good theory: Exploring theory based evaluation for comprehensive community for children and families. In J. P. Connell, A. C. Kubisch, L. B. Schorr, & C. H. Weiss (Eds.), *New approaches to evaluating community initiatives: Concepts, methods, and contexts* (pp. 65–92). Washington, DC: Aspen Institute.

World Bank. (2015). *World development report 2015: Mind, society and behavior*. Washington, DC: World Bank. doi:

10.1596/978-1-4648-0342-0.

APPENDIX 1: TOOL FOR IDENTIFYING POTENTIAL SUCCESS MEASURES

(Blank cells indicate where to put your answers to questions on the same row).

Theory of change component	SITUATION ASSESSMENT (CURRENT STATE)			INTERVENTION THEORY OF CHANGE (FUTURE STATE)		
	Ask yourself about the problem	Current State Indicators (narrative statements of current conditions for people or a thing)	Measures / Evidence (empirical measures of current conditions for people or a thing)	Ask yourself about the result you want	Future State Indicators (narrative statements of the intended change in conditions for people or a thing)	Measures (empirical measures of the intended change in conditions for people or a thing); possible KPIs / targets
Ultimate/End Outcomes/ Benefits Societal/ Cultural/ Economic/ Environment Current Situation This row defines your LAG indicators and measures – changes you expect following your interventions.	What is the ultimate problem? What is the risk to public health, the environment, public safety, the economy, etc.? What is the severity of the risk? How do our current efforts influence the ultimate outcomes/situation?			When will the problem/risk no longer be an issue? What will the problem/risk look like when it is no longer an issue? How do the interventions lineup with policy objectives? In what ways can we attribute changes, or demonstrate contribution, from our interventions?		

<p>Intermediate Outcomes/ Target Audience Behaviour</p> <p>This row defines your LEAD indicators and measures – behaviour changes you expect to directly result from your intervention.</p>	<p>What behaviours by whom are having a negative impact on the situation?</p> <p>What behaviours would you like to see change?</p>		<p>What behaviours by whom would have a positive impact on the situation?</p> <p>How do the behaviours need to change?</p> <p>Define precisely, and prioritise (i.e., see Chapter 5)</p>		
<p>Immediate Outcomes/ Behavioural Antecedents</p> <p>Target audience Knowledge, beliefs, attitudes, capabilities</p> <p>This row defines your ANTECEDENT indicators and measures – typically drivers and barriers you know or believe</p>	<p>What aspects of the target audience's knowledge, beliefs, attitudes, capabilities etc. are believed to be antecedents to <u>problematic behaviours</u>?</p> <p>How do these antecedents contribute to problematic behaviours?</p>		<p>What gaps exist in the target audience's knowledge, beliefs, attitudes, capabilities etc.? Which ones are believed to be antecedents to <u>desired behaviours</u>?</p> <p>How will the target audience's knowledge, beliefs, attitudes, capabilities etc. change?</p> <p>How do these antecedents contribute to desired behaviours?</p>		

lead to
behaviours.

<p>Reactions</p> <p>This row defines your REACTION indicators and measures – how people experience services, programs and interventions</p> <p>You may have relevant examples here from existing customer satisfaction, engagement and participation measures and feedback.</p>	<p>How do our target audiences currently experience existing services, programs etc., and how might that be contributing to problematic behaviours?</p>		<p>How do we expect our target audience(s) to react immediately to the deliverables?</p>	
--	---	--	--	--

<p>Reach</p> <p>Target Audience and Stakeholder Participation</p> <p>This row defines your PARTICIPATION indicators and measures – what proportion of your target audience access services, programs and interventions.</p> <p>It can also be important in identifying others, who themselves may need to act differently to help the target audience.</p>	<p>Are there gaps in the participation or engagement of groups which are critical to achieving your objective?</p> <p>(Defines target population(s) for behaviour change, and key stakeholders who may mediate, enable or exacerbate)</p> <p>If we don't have good access to the target audience, who does?</p>		<p>Who is this initiative intended to reach or who will be affected?</p> <p>Whose behaviour needs to change? (group/population)</p> <p>How will the participation or engagement of important groups change as a result of your program?</p> <p>With whom and how are we partnering to reach the target audiences?</p>		
--	---	--	---	--	--

<p>Outputs</p> <p>Programs/ supports in place for addressing the problem</p> <p>This row defines your OUTPUT indicators and measures – the ‘work’ of the initiative, often reporting on deliverables specified in project plans.</p>	<p>Are there gaps in the current suite of supports/programs/ services in place to address the problem or risk?</p> <p>What is current performance against time, quality and budget targets?</p>	<p>What program or service will we deliver in order to fill the gaps?</p> <p>What time, quality and budget targets must be met to deliver them successfully?</p>		
---	---	--	--	--

<p>Activities</p> <p>Internal practices</p> <p>Less often expressed as indicators and measures, but a chance to articulate foundational work required to produce outputs.</p>	<p>Are there problems with the current delivery practices?</p> <p>Are there programs or services being offered in other jurisdictions that are demonstrating better results than our programs or services?</p>	<p>What will we do differently?</p> <p>What actions or work will be done?</p> <p>What services (interventions) will be delivered?</p>		
--	--	---	--	--

Inputs	Are there gaps in the financial, human, or technical resources available?			What resources do we have for this initiative? What additional information would you like to know to improve delivery?		
Internal resources This row can provide indicators and measures for budget reporting, internal quality assessment, capability and capacity, as well as data quality.	Is data availability a problem? Are there gaps in our capabilities and capacities?			What internal capacities and capabilities are required to deliver improved activities?		

THE METHOD

CHAPTER 11 Intervention trials: How to know if your intervention works

By Wing Hsieh, Bernice Plant, Fraser Tull,
Celine Klemm and Fernanda Mata

www.behaviourworksaustralia.org



TRIAL

INTRODUCTION

Marius has started an exciting new career in youth justice. Having grown up in a “rough” neighbourhood and knowing people who have spent much of their life in prison, he desperately wants to make a difference. He knows that the problems are complex and efforts to create positive change will not be easy, but he is armed with resources (although limited) and keen to act. As it so happens, soon after starting his new position, he comes across an HBO documentary series called Beyond Scared Straight. Based in the US, each episode follows a small group of at-risk teens as they visit a prison and get to see what life is really like behind bars. Realising the show is intentionally designed for shock-value and entertainment (and not necessarily based on real initiatives) he starts looking into similar programs across the world. He discovers a range of approaches, but all having the same quality of giving at-risk teens the opportunity to hear the stories and advice from inmates, and from employees working within the prison. He finds numerous examples of these programs advertised on prison websites along with quotes from teens who had successfully gone through the program. Stories of appreciation to the program for “getting me back on track” came through and the overwhelming conclusion was that these programs work! Inspired by these findings, Marius starts thinking about how he could implement such a program in his jurisdiction. But if he were to implement the program, how would he know if it worked?

In previous chapters, you learnt how to develop behavioural interventions and identify success measures. However, even when an intervention is based on established theory and evidence, there is no guarantee that the intervention will work. In this chapter, we show you how to reliably evaluate whether your intervention will have a positive impact on your success measures. Only with a strong evaluation framework can you be confident in knowing the actual effectiveness of your intervention. Now, Marius may not have read our previous chapters and instead relied on his intuition of “what works”, but as you’ll come to understand, he also made the mistake of relying on weak evidence to support his decision making.¹¹ Quotes from a few satisfied participants and unsubstantiated claims made by program administrators is not sufficient for determining the effectiveness of an intervention, but can be helpful to understand the acceptability of an intervention. To determine effectiveness, an evaluation needs to be undertaken.

In this chapter, we introduce a suite of common approaches for evaluation design, describe some key challenges for evaluating whether something has worked in the real world, and discuss several considerations when deciding what evaluation approach to use.

COMMON EVALUATION APPROACHES

Whether we are trialling or delivering our intervention, a robust evaluation design is needed to detect a change in success measures and to ensure that outcomes are attributable to the intervention as much as possible, and that there are no alternative explanations to the changes we observe in our outcome, other than the intervention itself. In the following section, we will discuss three types of evaluation designs that can be used for effective intervention evaluation, and help you pick a design that is right for you. We have chosen these designs because we have found them to be the most useful in our work with policy and program partners. There are many more designs available, and as you delve further into

¹¹ For more information on the evidence hierarchy, see Petrisor and Bhandari (2007).

evaluation and trial design, you may wish to refer to other resources to explore other approaches (e.g., Mathews & Simpson, 2020). There are also qualitative evaluation

designs which we do not consider here as they tend to relate to evaluation questions other than effectiveness (e.g., when interested in understanding whether an intervention is *acceptable* to your target audience).

Randomised Control Trial (RCT)

Case study: Victorian Managed Insurance Authority (VMIA)

Health service boards have a critical role to play in overseeing hospitals and health services. Further, individual board members are responsible for navigating a range of complex health and safety issues to ensure patient safety. Thus, it is essential they have the skills and confidence required to effectively share information, communicate concerns and engage in difficult conversations when needed.

For those new to the board setting, there are relatively few opportunities to practice and develop these skills in a controlled environment. As part of the VMIA Patient Safety Research and Innovation Program,

BehaviourWorks Australia partnered with VMIA to understand if simulation training (the intervention) with an actor could help. We conducted a randomised control trial to assess the effectiveness of simulation-based training (compared with no training) in improving board members' skills and confidence in communicating effectively in board meetings.

Randomised control trials are regarded as the gold standard for examining the effectiveness of behavioural interventions and establishing causal conclusions (Gerber & Green, 2012). This design involves the random allocation of participants or groups of participants to two or more separate groups (e.g., intervention group and control group) (Gerber & Green, 2012). The intervention group will receive the intervention that is being tested while the control will typically receive a business-as-usual program or nothing at all. This approach provides a tool to investigate cause-effect relationships between an intervention and outcome. This is because randomisation balances participant characteristics between the groups, which allows attribution of any differences in outcome measures to the intervention. This design provides the strongest evidence for drawing cause-effect conclusions between the intervention and the outcomes of interest.

In the VMIA case study, twelve health services boards were randomly allocated to intervention or control groups. We randomly allocated boards, taking into account whether they were regional or metropolitan (stratified randomisation) to ensure that comparable numbers of regional and metropolitan boards were allocated to each group of the trial. The intervention, involving immersive, simulation-based training of health services, was delivered to boards allocated to the intervention group. The outcomes of interest in this intervention (perceived skills and confidence in boardroom communication) were recorded for both control and intervention groups immediately before the intervention was delivered to the

intervention group (Time 1), as well as three months post-intervention (Time 2). Figure 1 illustrates the design of a randomised control trial.

RANDOMISED CONTROL TRIAL STUDY DESIGN

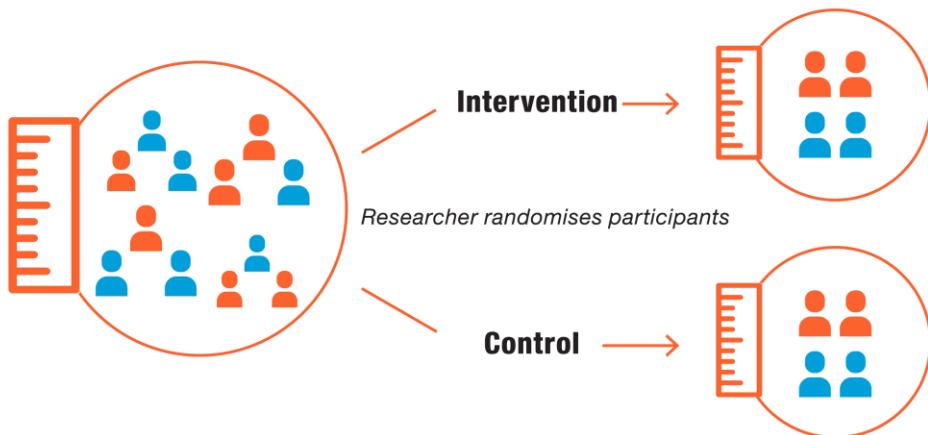


Figure 1: Illustration of a randomised control study design

Three months after the intervention was delivered (Time 2), health board members who were allocated to the intervention group reported significant improvements in their skills and confidence in communicating effectively during meetings, compared to the control group. The key characteristics of randomised control trials, such as randomisation and inclusion of a control group, reassured the research team and VMIA that the improvements observed in the outcome variables (board members' skills and confidence) were attributable to the intervention delivered (simulation-based training). If you want to know more about what we found, visit the [BWA website](#) or access this [publication](#).

Not all control groups are created equal

What makes a suitable control condition? The nature of your control condition will depend to some extent on your research question; however, there are some general approaches, each with their strengths and weaknesses, which could help you to navigate, select or design a suitable comparison to your intervention.

No-treatment: In a “no-treatment” condition, people are not exposed to any intervention or treatment – in a sense, you can think of this as a “do nothing”. While comparing your intervention to “no intervention” is better than not including any control group, relatively speaking, this approach is often considered less rigorous as a control condition. This is because while you might be able to demonstrate that your intervention has an effect, you might be left wondering if this was because your participants were simply doing “something” (anything!) rather than “doing nothing”. However, in some instances, this is the most feasible comparison group.

Current practice: This is typically considered a more rigorous approach than a no-treatment control because you are testing the effects of your intervention in comparison to “doing something else” (depending on the nature of the current practice, this may also be an

“active control” condition). This is a common approach we take with our projects at BWA, as we often want to test whether there is value in applying behavioural insights to existing material used by our partner organisations.

Intervention, minus its active ingredient: In some cases, you may prefer or may need to demonstrate that your intervention works because of a particular “ingredient” or via a proposed mechanism. In such cases, you may choose to compare your intervention to a condition that is essentially the same, except that it does not contain the specific component believed to be the driving force behind your intervention’s effects. When this approach is taken, an additional no-treatment or current practice control condition(s) may be tested concurrently – this is because if no difference was found between your intervention and control (minus active ingredient) conditions, an additional control would help to clarify whether both intervention conditions had an impact or neither of them did.

While the strength of RCTs is in being able to provide the strongest evidence of causation between a treatment and outcomes, it is not always feasible to randomly allocate participants into a treatment and control group. For example, if you wanted to understand whether an information poster was effective, you might put up the poster only on randomly selected floors of a building but there is a high chance that staff might move across floors, and those on the “control” floors might end up seeing the poster on “intervention” floors. In this instance there is a risk of cross-contamination between the intervention and control groups. An alternative evaluation approach that does not require randomisation could therefore be considered, such as the difference-in-differences approach.

Pre-Post intervention design with Control Group (Quasi-experimental “Difference-indifferences approach”)

Case study: VicRoads

Each year, the Victorian government agency VicRoads, which provides driver licensing and vehicle registrations services, processes over 900,000 change of address requests. In response to this, VicRoads decided to implement an online self-service tool that makes it possible for customers to change their address online. However, a key risk for the success of these types of initiatives is a low uptake of these services by customers.

In 2017, VicRoads commissioned BehaviourWorks Australia for a project aimed at increasing the use of their e-government services. Together with VicRoads employees, we developed an intervention drawing on behavioural insights alongside a trial to evaluate its effectiveness.

In both the intervention and control areas, participants visiting a customer service centre were greeted by a concierge and asked about the reason for the visit. In the control group, customers who said they were only visiting the centre to change their address were given a ticket and directed to wait until a staff member called their number. This was VicRoads default practice at the time of the trial. In the trial area, the developed intervention was tested.

In the VicRoads example, we chose a pre-post intervention design with a control group, also called difference-indifferences approach. Using a randomised control trial design was not suitable because of a high risk of crosscontamination – the risk that participants had moved between the intervention and control area which meant that some of the participants in the control group could have learned about the online self-service tool. Instead, a prepost design with a control group was used, as that design helps to overcome the problems associated with crosscontamination.

A pre-post design with a control group is a type of quasi-experimental design, in which participants are not randomly allocated to conditions. Instead, intervention and control groups are purposely selected to ensure that they are similar to each other. This is important because if the groups are similar to each, then the differences in the outcomes of interest are likely to be due to the intervention. The impact of the intervention is measured at two points in time - before and after the intervention – for both the intervention and the control group (Figure 2).

PRE-POST WITH CONTROL STUDY DESIGN

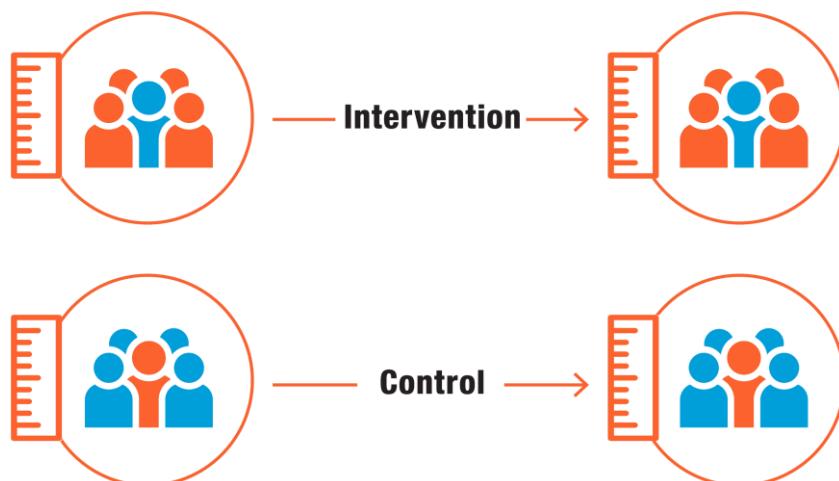


Figure 2: Illustration of a pre-post intervention design with control group (also called “difference-in-differences approach”)

As part of the intervention, customers who said their sole reason to attend the customer centre was to change their address were invited to complete the transaction online using one of two in-store tablet devices, rather than taking a ticket and waiting in a queue, and were offered support from on-site staff if needed. For the purpose of evaluation, customers in the intervention area who attempted to use the tablet to change their address online were invited by VicRoads staff to respond to a very short survey after completing the transaction. The survey measured if participants changed their address online or in another way (e.g., by phone, letter, in person), satisfaction with the service and what they liked and disliked about using the tablet for the service.

In a pre-post design with a control group, the effectiveness of the intervention is evaluated by comparing the change in outcomes before and after the intervention among the group that received the intervention and a control group that did not receive it over the same period of time (Imbens & Wooldridge, 2009; Lechner, 2011). Simply put, the effect is estimated by calculating the change that was observed in the intervention group and subtracting from it

the change that occurred without an intervention (in the control group) (Fredriksson & Oliveira, 2019). The assumption is that if there had not been an intervention, the treatment group would have followed the same trend (e.g., in uptake of the self-service online tool) as the control group (Gertler, Martinez, Premand, Rawlings, & Vermeersch, 2016).

For the VicRoads study, we compared the percentage of transactions completed online in the two months before and two months after the intervention commenced. The study showed that the intervention increased customers' use of the online service and that customers indicated that they appreciated the time-saving benefits, ease of use and customer support. If you want to know more about what we found, visit the [BWA website](#) or access this [publication](#).

This evaluation approach is well-suited to evaluate the effect of introducing new legislation, policies, programs or reform. For example, if a country were to introduce a school reform, it is likely not feasible to randomly select certain schools for an intervention (i.e., they receive the reform) and others not. In this case, the reform could be rolled out in a staged manner, state-by-state. A different state in which the reform has (not yet) been implemented can be used as a comparison and control group using a pre-post design with a control group (Schwerdt & Woessmann, 2020).

It would have been tempting to evaluate the impact of the intervention by using a pre-post design, without a control group (more on this design in the next section), given that we had data from both before and after the intervention started. This would have meant simply comparing the percentage of participants using the online service in the trial areas before and after the intervention was implemented. This type of design, however, would not have allowed us to attribute changes in the uptake of the online self-service tool to the intervention. Observed changes could also have been caused by other factors, for example, a general, perhaps natural, trend in the population to use more online services or increased digital literacy that occurred regardless of the e-government intervention. Where it is possible, it is preferred to have a control group for comparison. However, there may be circumstances where it is not possible to have a control group. For example, because of equity concerns, it is necessary to open eligibility to a program to everyone. In this instance, it may be necessary to consider a prepost intervention design without a control group.

Pre-post intervention design without a Control group (Quasiexperimental)

Case study: Victorian Department of Health and Human Services (DHHS)

In 2015, the Victorian Department of Health and Human Services (DHHS) conducted a major review of Victoria's ambulance services. Among the issues identified was the fact that there had been a huge rise in the number of non-emergency calls to Triple Zero and that these calls were diverting up to 10 ambulances a day away from urgent cases.

In response to this, DHHS partnered with BehaviourWorks Australia and The Shannon Company to improve public awareness of the role of ambulance services, including when it is appropriate to dial Triple Zero. We developed a mass media communication campaign

to address this challenge and to illustrate the types of medical emergencies requiring an ambulance. Subsequently, we conducted a pre-post intervention trial without a control group (obtaining a comparison group was not feasible) to evaluate the effectiveness of the campaign in shifting attitudes towards using ambulances.

Pre-post intervention designs (without control), another type of quasi-experimental study design, are also commonly used to evaluate the effectiveness of behavioural interventions. As shown in Figure 3, this design involves measurements for a single group taken both before and after an intervention is delivered, and is often used in cases where a control group cannot be recruited or where randomisation into treatment and control groups is difficult. If the post-intervention scores are better than the pre-intervention scores, then it may be reasonable to assume that the intervention was responsible for the improvement. However, there may be other explanations for why the post-intervention scores have improved (e.g., environmental factors). Therefore, we cannot be certain that the changes observed in the outcome measure are the direct result of the intervention examined (Miller, Smith and Pugatch, 2020).

PRE-POST STUDY DESIGN



Figure 3: Illustration of a pre-post intervention design without control group

In the DHHS example, we measured behavioural intentions and attitudes towards ambulances before the campaign was launched (pre-intervention) and after the campaign had been active for four months and ten months (post-intervention). Further, actual behaviour was measured using data for calls from Triple Zero at three different time points within a two-year time frame. We looked at data (consisting of both high and low priority calls) both during the evaluation phase of the research (post-intervention), as well as three years prior, to provide a pre-intervention (baseline) comparison.

In addition to improvements in attitudes towards the appropriate use of Triple Zero, the evaluation results showed a significant reduction in the use of Triple Zero after the behavioural campaign was initiated, compared to the same time the previous year. While we cannot rule out the influence of external factors that may have contributed to these findings (due to the absence of a control group), our findings suggest that our evidencebased campaign was successful in shifting attitudes towards ambulance services.

Pre-post intervention designs could be improved by ongoing monitoring and evaluation of the intervention being delivered. In other words, rather than just two instances of

measurement only (before and after), the effectiveness of the intervention can also be examined one month later, and/or two months later, etc., after the intervention has been delivered. Repeated measures before and after an intervention is delivered can help us identify other factors that may have influenced the outcomes as well as improve our understanding of sustained and long-term behaviour change. What we do not want to see are immediate effects only and waning effects after that.

HOW DO YOU DECIDE WHICH EVALUATION APPROACH IS RIGHT FOR YOU?

In the previous section, we guided you through three common evaluation approaches that help with establishing “What works?”. Now you may be wondering how to choose an approach that meets your needs, and so this section aims to highlight some key considerations when selecting your evaluation approach.

Criteria for selecting an evaluation approach

When it comes to selecting an evaluation approach, there are three criteria to consider:

1. **Evaluation question** - What do you want to answer with the evaluation?
2. **Budget** - Does the approach fit within your resourcing budget?
3. **Feasibility** - Is the approach practical or feasible?

Different approaches have different strengths and weaknesses. For example, a randomised control trial is the best approach to answer the question of causation (i.e., “Did the intervention cause a change in the outcome?”). We summarised the strengths and weaknesses in Table 1. However, it might not be feasible to randomise – for instance, it is hard to prevent customers from finding out about an online form as in the VicRoads example.

Table 1: Summary of selected strengths and weaknesses of common evaluation approaches

Approach	Strength	Weakness	Best used when...
Randomised control trial (randomly assigning members of your sample to different study conditions and assessing which study condition performed better)	Best design for attributing causation from the intervention to the outcomes. Reduces allocation and selection biases.	Practical challenges with random allocation (e.g., crosscontamination).	You need strong evidence of whether the intervention causes the target outcome and random allocation is feasible.

Pre-post with control (one group of your sample receives the intervention while another group does not receive the intervention, and you assess whether there was greater improvement among the group who receive the intervention)	Avoids practical challenges with random allocation. Takes into consideration changes that may occur due to other factors, besides the intervention. Can estimate a causal effect using observational data if assumptions are met.	Lack of randomisation means that the results may be due to a particular characteristic that drives the outcome and is more heavily represented in one of the groups. Cannot be used if the composition of groups before and after an intervention is likely to change.	You want some evidence of whether an intervention worked, a control group is possible to define but random allocation between treatment and control groups is not feasible.
Pre-Post without control (everyone in your sample received the intervention and you assess whether outcomes are better after the intervention, compared to before the intervention, across your sample)	Simple to implement	The absence of a control group makes it difficult to know whether the results are due to external or environmental factors. Results are only suggestive of impact.	You want some evidence of whether an intervention worked, but a control / comparison group is not feasible.

However, as illustrated throughout the “Common evaluation approaches” section, the trial design you select will not only depend on your evaluation questions but will be influenced or constrained by a range of additional factors, including feasibility (e.g., of randomisation), resourcing (e.g., available budget, time, sample, team capability) and ethical considerations. Below we explore some of the key considerations. Ultimately, the evaluation approach chosen will weigh up these factors.

Different evaluation questions: “What works?”, “How?” or “When?”

Behaviour-change practitioners are typically most interested in understanding the impact of an intervention on their outcome(s) of interest (or “*What works?*”) because these factors are often considered more important for key stakeholders (compared to, for example, contributing towards theory, which might be more important for academics). However, there are occasions where you might want to understand the process underlying *how* your intervention works or to gain an understanding of whether your intervention works under particular circumstances or boundary conditions (*when* your intervention works).

For example, let’s imagine that you wish to demonstrate your intervention (e.g., a persuasive video) works via a particular psychological process (e.g., increased self-efficacy) or other theoretical mechanism of interest. In such cases, it might be possible to measure the impact of the intervention on the construct of interest (e.g., self-efficacy) prior to measuring its impact on your outcome of interest. However, a challenge with this approach is that the additional measurement could potentially contaminate the effects of the intervention on your outcome, and so you risk interfering with your impact assessment. Asking people to indicate their level of self-efficacy after watching a persuasive video could result in them reflecting on the video and the construct more than they would otherwise. If you subsequently find that

your video is impactful, you might be left wondering whether it was the video or the fact that you asked participants to reflect on their self-efficacy. This is a likely reason that potential mechanisms are relatively rarely explored in the field.

Of course, as was raised in relation to success measures (see Chapter 10), neglecting to explore “how” an intervention works during an evaluation can have implications for refining and developing effective interventions, as well as for making theoretical contributions within the field (also see discussions about process evaluations in Abraham & Denford, 2020). However, if testing “how” your intervention works is a priority, there are alternative ways to gain insights into potential underlying processes, including: experimental approaches, such as manipulating components of your intervention in a systematic way; testing theoretical mechanisms via subsequent research, which could include manipulating the causal direction (you can read more about causal order in Hayes, 2018).

Available budget for intervention evaluation

The available budget for a project can influence the nature of your evaluation approach. From our experience of discussions with participants during our [training sessions](#), it appears to be a commonly-held belief that RCTs are more expensive to run, and this can deter people from considering them. We recommend you do not rule out RCTs on the basis of cost alone, as an RCT is not necessarily more expensive to run. For instance, since RCTs may only require a single measurement from each participant, the duration of your sessions may be shorter for each participant than designs requiring repeated measurements, thereby reducing costs. Particularly when you have relatively inexpensive methods of delivering your intervention and measuring its impact, such as [using SMS to increase vaccination rates for the HPV vaccine](#), an RCT does not need to be costly.

Feasibility from the perspective of your organisation and/or the trial site

As mentioned earlier in this chapter, if potential participants work in close proximity to each other, it may be more likely that intervention spillover effects or contamination will occur, reducing the feasibility of conducting a highquality RCT. For example, randomising individual participants to either an intervention or control condition may not be feasible in a hospital setting where clinicians work closely together and may either share or discuss the intervention with each other. In such a case, a more appropriate approach might be to allocate, say, one ward to serve as an “intervention” group and another ward to serve as a “control” group (pre-post with a control group).

Potential sample size

Sometimes, your “pool” of potential participants could be known to be small, or you might experience challenges with reaching, recruiting, and/or retaining a sufficient number of participants that would typically allow a comparison between two or more groups (e.g., control versus intervention groups). In such cases, it could be safer to consider designs that provide opportunities that minimise the possibility that any observed effect is due to factors other than the intervention (e.g., pre-post with a control group) or designs that do not require betweensubjects comparisons (e.g., pre-post without a control group). In behaviour change, and particularly for real world trials, often the sample size is determined on a pragmatic basis (e.g., considering time and budget available). You may be wondering, then, how we were able to run an RCT with a sample size of 57 participants for the VMIA case study

described earlier in this chapter? This is because of a number of additional factors that influence what sample size may be required, which are beyond the scope of this chapter – but for those who are interested, you can read the ‘Sample size calculation’ section of the associated [publication](#). If you have the opportunity, we would recommend calculating the sample size that would be required to detect an effect beforehand, which can help inform the design(s) that might be more appropriate for your situation. You may wish to consider online sample size calculators such as [G*Power](#) or [GLIMMPS](#).

Ethical considerations

In addition to a range of practical considerations, it is important to consider potential ethical issues that may arise when selecting the evaluation approach that is right for you. For example, if it is not ethical to collect data before the trial commences (baseline data) as part of your evaluation, pre-post designs would not be appropriate, thereby shifting your focus to other evaluation approaches. In addition, ethical issues can arise when it comes to using control groups. For example, in the case of RCTs, it may not be ethical to randomly assign people to different interventions or to delay intervention (i.e., for the control group). If possible, participants assigned to a control condition should have the opportunity to access the intervention at the conclusion of the study. These are just a few examples of ethical considerations that could influence your choice of evaluation approach (as a research unit of an Australian university, all of our methods must align with the [National Statement on Ethical conduct in Human Research](#)). For an example of broader discussions regarding the ethics of behaviour change, see [The Behavioural Scientist's Ethical Checklist](#).

Linking to outcomes and measures

While the above considerations have been discussed in relation to trial design, you will also need to consider the outcomes and measures of interest alongside your trial design (see Chapter 10). For instance, some combinations of recruitment methods and outcome measures will put a larger dent in your budget than others (e.g., recruiting and offering an incentive to participants via a panel company for individual 30-minute interviews versus recruiting participants via social media for an online survey, where the incentive takes the form of a prize draw). While success indicators may be selected prior to designing your trial, your selection of measures has an influence on, and may constrain, the trial design and vice versa. Consequently, it is advisable to consider these concurrently, or at least you should keep trial design in mind as you explore options for your measures, so that you do not limit your options!

Online or laboratory studies

While we have focused on testing in the real world in this chapter, there will be situations where testing in an online or in a “lab” setting would be suitable. Again, there are several factors that will likely influence whether you test or evaluate your intervention in a real-world setting, ranging from whether it is the preferred approach (in line with your evaluation question) to whether it is feasible (in line with your practical considerations and constraints). Often when conducting trials there is a balancing act between having confidence that any effect you have observed can be attributed to your intervention (*internal validity*) and confidence that your effect or findings can be applied to other situations or groups (*external validity*), including how likely they are to translate to the “real world” or “field” (*ecological validity*). Generally speaking, trials conducted in the real-world tend to be higher in external validity and lab-based experiments tend to be higher in internal validity. What we want to

emphasise at this point, is that conducting trials online or in “lab” settings are not inherently “bad” – they can be valuable and even superior to trials conducted in more ecologically valid settings.

Further, there may be creative ways for you to enhance the ecological validity of a trial conducted online or in the lab. For instance, this could be in the form of using or designing creative success measures or could take the form of assessing the impact of your intervention via a creative task that serves as a proxy for what might occur in the real world. An example in the case of measurement during an online trial might be that instead of asking people to indicate via a questionnaire how strong their intention is (e.g., to quit smoking), you might present information about how to access or engage with a service that could help them to quit smoking and measure the amount of time participants spend on the page (i.e., as a proxy for their intention to quit smoking).

The following case study aims to demonstrate a creative way of testing the effects of an intervention during an online trial. In this case study, we attempted to increase the ecological validity of the task and the way that participants could respond during the task.

Case study: An example of enhanced ecological validity for an online trial

The trial purpose: The [VMIA inter-hospital transfers trial](#) tested the effectiveness and acceptability of a video that reminded clinicians who commonly receive transfer requests from a “sending” hospital that clinical staff in other hospitals and areas across the health system share the same goals.

The challenge we faced: The behaviour of interest was related to conversations that are held between clinicians from “sending” and “receiving” hospitals. For ethical and practical reasons, it was not feasible for us to test the effects of the video by observing or recording real phone conversations between clinicians in hospitals during non-critical inter-hospital transfer requests.

How we addressed this challenge: We compared outcomes before and after exposure to the video using audio vignettes presented during an online study to simulate transfer request phone calls from smaller “sending” hospitals for non-critical patients. That is, participating clinicians were prompted at relevant moments during the trial to click and listen to an audio recording which simulated a caller’s request. Of course, audio vignettes are not equivalent to dynamic conversations over the phone; however, in combination with our success measures, this was a creative solution that allowed us to examine the effects of our intervention via an online study in a more ecologically valid manner.

Weighing up evaluation approaches

The factors covered in this section cannot really be considered in isolation. For instance, there may be multiple ways that you could recruit members of your target population, which has implications for your budget, your potential sample size, and the potential ethical issues you will need to consider. Similarly, the potential time frame you have available for your trial will influence the overall design, but also the potential measures you use – for example,

shorter project timeframes tend to push designs towards examining immediate effects only, using lower-risk opt-in forms of recruitment and measures lower in ecological validity (versus covert observation of observable behaviour in the form of a field trial) – again, these factors need to be considered together.

In many respects, to go from a long list of potential trial design options to a short list could be achieved by applying a prioritisation process, such as that described in the context of intervention design (see Chapter 8; also see Chapter 5 on prioritisation in the context of selecting target behaviours) could be applied to trial design. Again, extensive consultation with key stakeholders is crucial for weighting the feasibility criteria – this may include input from your partner organisation, your team, trial site (if applicable), the target audience, and/or the ethics office. However, we typically do not apply a formal prioritisation process for trial design. Instead, we might draft up 1–3 trial design options and discuss their advantages and disadvantages with our partner organisation and other key stakeholders as a way of ensuring there is a shared understanding about the approaches and priorities before selecting and developing the final trial design.

Given the number of factors that need to be considered and that often multiple stakeholders are involved, selecting an appropriate trial design can be a complex and challenging process. There is always some trade-off (often between internal vs. external validity) -- there is no “perfect” design that will be able to do *everything*. In a sense, the “right” design for you is one that aligns with your priorities – if executed well, a lower-quality design may be superior in meeting your needs to a high-quality design executed poorly.

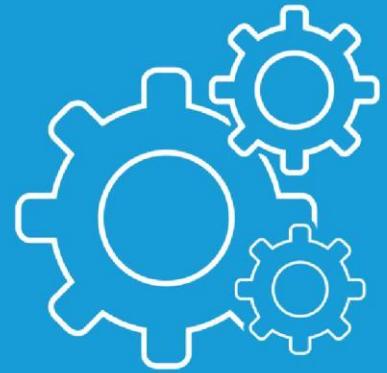
CONCLUSION

In this chapter, we have introduced the importance of trialling interventions, the common evaluation approaches that are used to understand whether an intervention “worked” and some of the key considerations for designing reliable evaluation frameworks. There are, of course, many more study designs and approaches than we have covered in this chapter. Take for example Marius’ experiences in trying to determine the effectiveness of juvenile awareness programs. His early sources were essentially one person’s opinion about the programs. Fortunately for Marius, researchers have already undertaken a thorough evaluation and demonstrated that these juvenile awareness programs have no impact on the future behaviour of at-risk teens (van der Put, Boekhout van Solinge, Stams, Hoeve, & Assink, 2021). This overview of the results of many randomised control trials and quasiexperimental designs provided a very robust evaluation of these programs. Ultimately though, study design should depend on what evaluation question you want to answer, the budget you have available and the feasibility of your options. And of course, once you have determined that you have a successful intervention, you may wish to scale up the intervention to more sites or across different groups. It is to this issue of scaling that we turn to next, in Chapter 11.

REFERENCES / FURTHER READING

- Abraham, C., & Denford, S. (2020). Design, implementation, and evaluation of behavior change interventions: A ten-task guide. In M. S. Hagger, L. D. Cameron, K. Hamilton, N. Hankonen & T. Lintunen (Eds.), *The handbook of behavior change* (pp. 269-284). Cambridge, UK: Cambridge University Press.
- BehaviourWorks Australia. (n.d.). Communication in the boardroom. Retrieved 15 October 2021 from: <https://www.behaviourworksaustralia.org/victorian-managed-insuranceauthority/trial-7-communicationin-the-boardroom/>
- BehaviourWorks Australia. (2017). Using SMS to increase vaccination rates among secondary school kids. Retrieved 15 October 2021 from: <https://www.behaviourworksaustralia.org/portfolios/using-smsmessaging-increasevaccination-rates/>
- BehaviourWorks Australia. (2017). Increasing digital uptake of government services. Retrieved 15 October 2021 from: <https://www.behaviourworksaustralia.org/portfolios/increasing-digital-takegovernment/>
- BehaviourWorks Australia. (2020). Patient Safety Research and Innovation Program Retrieved 15 October 2021 from: <https://www.behaviourworksaustralia.org/victorianmanaged-insurance-authority/>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.
(G*Power calculator) Retrieved 15 October 2021 from: <https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-undarbeitspsychologie/gpower>
- Faulkner, N., Jorgensen, B., & Koufariotis, G. (2019). Can behavioural interventions increase citizens' use of egovernment? Evidence from a quasi-experimental trial. *Government Information Quarterly*, 36(1), 61–68. doi:10.1016/j.giq.2018.10.009
- Faulkner, N., Wright, B., Lennox, A., Bismark, M., Boag, J., Boffa, S., ... & Bragge, P. (2020). Simulation-based training for increasing health service board members' effectiveness: A cluster randomised controlled trial. *BMJ Open*, 10(12), e034994. doi:10.1136/bmjopen-2019-034994
- Fredriksson, A., & Oliveira, G. M. D. (2019). Impact evaluation using difference-indifferences. *RAUSP Management Journal*, 54, 519-532. doi:10.1108/RAUSP-052019-0112
- General Linear Mixed Model Power and Sample Size (GLIMMPS). Retrieved 15 October 2021 from: <https://glimmpse.samplesizeshop.org/>

- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016). *Impact evaluation in practice* (2nd ed.). Washington, DC: Inter-American Development Bank and World Bank.
- Gerber, A. S. & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. New York, NY: W. W. Norton & Company.
- Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: a regression-based approach* (2nd ed.). New York, NY: Guilford Press.
- Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation.
Journal of Economic Literature, 47(1), 5-86. doi:10.1257/jel.47.1.5
- Jachimowicz, J., Matz, S., & Polonski, V. (2017). The behavioral scientist's ethics checklist. *Behavioral Scientist*. Retrieved 18 October 2021 from:
<https://behavioralscientist.org/behavioral-scientists-ethics-checklist/>
- Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends in Econometrics*, 4(3), 165-224. doi:10.1561/0800000014
- Matthews, L. & Simpson, S. A. (2020). Evaluation of behavior change interventions. In M. S. Hagger, L. D. Cameron, K. Hamilton, N. Hankonen & T. Lintunen (Eds.), *The handbook of behavior change* (pp. 318332). Cambridge, UK: Cambridge University Press.
- Miller, C. J., Smith, S. N., Pugatch, M (2020). Experimental and quasi-experimental designs in implementation research. *Psychiatry Research*, 283, 112452. doi:10.1016/j.psychres.2019.06.027
- The National Health and Medical Research Council, the Australian Research Council and Universities Australia (2018). The National Statement on Ethical Conduct in Human Research 2007 (Updated 2018). Retrieved 17 October 2021 from:
https://www.nhmrc.gov.au/about-us/publications/national-statementethical-conducthuman-research-2007-updated-2018#toc_15
- Petrisor, B., & Bhandari, M. (2007). The hierarchy of evidence: Levels and grades of recommendation. *Indian Journal of Orthopaedics*, 41(1), 11–15. doi:10.4103/00195413.30519
- Schwerdt, G., & Woessmann, L. (2020). Empirical methods in the economics of education. In S. Bradley & C. Green (Eds.), *The economics of education* (2nd ed.) (pp. 3-20). Cambridge, MA: Academic Press.
- van der Put, C., Boekhout van Solinge, N., Stams, G., Hoeve, M., & Assink, M. (2021). Effects of awareness programs on juvenile delinquency: A three-level meta-analysis. *International Journal of Offender Therapy and Comparative Criminology*, 65(1), 68-91. doi:10.1177/0306624X20909239

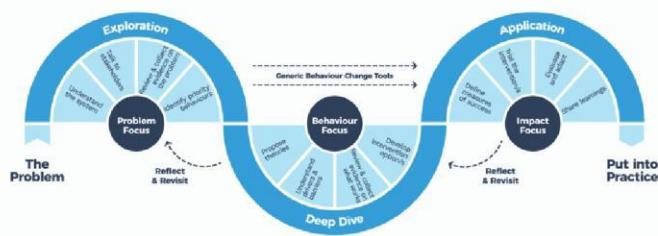


THE METHOD

CHAPTER 12: Making the most of an effective intervention

By Wing Hsieh, Peter Slattery, Stefan Kaufman,
Denise Goodwin, Alexander Saeri, Morgan Tear

www.behaviourworksaustralia.org



INTRODUCTION

Port Phillip Bay is one of Victoria's best-known waterways, with over 130 beaches that offer numerous recreational opportunities for swimmers and beachgoers. However, the result of increasing urbanisation surrounding the public use of the waterway is that there are occasions, typically following heavy rainfall, where swimmers are at risk due to pollution that washes into the Bay. The risk is particularly high when there is heavy rainfall one day, followed by high temperatures and sunshine the next. If people ingest the polluted water, they could become ill, most commonly with gastroenteritis (gastro), which can disproportionately affect people with weaker immune systems like children.

To reduce this risk, the Environment Protection Authority (EPA) worked in partnership with BehaviourWorks Australia (BWA), Federation University and Life Saving Victoria on a project focusing on developing interventions to prevent members of the public swimming on poor water quality days. The resulting interventions, which involved a combination of new signage and messages to mobile phones, were designed to inform beachgoers on high-risk days. They were found to be effective in trials, but was this enough? How did the EPA ensure that the successful trials led to broader behaviour change impacts, in this case, reducing the number of people swimming in the Bay on poor water quality days?

Previously in Chapter 11, we discussed a number of considerations when selecting evaluation strategies, introduced a suite of common evaluation designs, and reflected on some key challenges for evaluating whether something has worked in the real world. But once your trial has identified an effective intervention, you cannot assume that it will later be implemented in a way that allows it to have an impact. Too often, researchers and practitioners alike assume that as long as we find out what works, then that suffices to have impact in the real world. For example, Antman et al. (1992) found that it took 13 years after the publication of supportive evidence before medical experts began to recommend a new drug. During this time, numerous patients may have received suboptimal medical treatment - a prescription for a drug that was not the most effective one available. Many of these patients may therefore have suffered significantly, and some may even have died before their time. This is not an isolated case. It can take up to 17 years for research evidence to disseminate into health care practice (Balas and Boren, 2000; Morris et al., 2011) and perhaps only 14 per cent of available evidence enters daily clinical practice (Westfall et al., 2007). Furthermore, most social enterprises and, indeed, academic researchers, overinvest in the development and testing of novel innovations at the expense of real impact through the scale up of existing innovations (Seelos & Mair, 2017).

Reassuringly, there are several strategies that could be deliberately used to increase the likelihood that an effective intervention does result in real-world impact.¹² In this chapter, we explore three strategies to ensure that an effective intervention does lead to impact. Each pathway can increase your impact - i.e., desired behaviour and societal change - but each strategy approaches this goal from different directions and with an emphasis on different

¹² There is a whole science dedicated to this - "implementation science". Implementation science, founded in clinical medicine but now applied more broadly, is the study of methods and strategies to promote the uptake of evidence-informed practices into business as usual. It seeks to understand how to increase the impact of innovations, such as scaling, disseminating and translating new and effective interventions into practice.

activities and outputs. We will introduce you to the three strategies before providing guidance on how to apply each.

STRATEGIES FOR AMPLIFYING THE IMPACT OF EFFECTIVE INTERVENTIONS

In order for an effective intervention to have the most impact possible, it is important that it is able to be implemented beyond the initial pilot group, that the relevant decision-makers know about the intervention, and that the research is accessible to the relevant decision-makers. There are three strategies that help to achieve those goals - scaling, dissemination and knowledge translation/brokering (see Box A). In practice, you will likely need to use a combination of all three strategies, so it is helpful to understand when to use each one.

Box A: Three strategies for ensuring effective interventions have impact

Scaling is about taking an effective intervention and applying it to change the behaviour of more people. Its focus is mostly on the intervention itself, and in supporting organisations and communities to embed the intervention in their business as usual or increase the geographic reach of an intervention.

Dissemination is about broad, often public, communication of new behavioural science knowledge, which may or may not include specific behaviour change interventions, to audiences that may adapt, combine, or remix for their specific needs.

Knowledge translation is about increasing the awareness and use of insights and evidence between research, policy and practice. Knowledge translation typically involves *knowledge brokers* and can include activities such as vocational training and formal partnerships, or informal activities such as “watercooler” conversations about everyday challenges that connect ideas and evidence that would not otherwise meet.

There are two dimensions to consider when selecting a strategy: environment stability and the relationship strength. Figure 1 summarises when to choose which strategy. High **relationship strength** are contexts where there could be longstanding relationships of trust with stakeholders and a strong shared understanding of the “operating model” of how behavioural science projects work. Relationship strength is a key factor in choosing what type of strategy is feasible. Knowledge translation works best as a strategy when there is a strong relationship between knowledge brokers and practitioners/policymakers. This is because in such an environment, a knowledge broker has the connection to go straight to those who have the authority and responsibility to make the necessary decisions to

implement changes. In contrast, in situations of low relationship strength, dissemination is the best strategy as it does not rely on existing relationships.

High **environmental stability** are contexts where the target behaviour and/or audience is likely to stay similar over time and there is lots of scope to expand the reach of an effective intervention. Environmental stability matters particularly for scaling because for an intervention to apply to many different contexts, there needs to be a substantial degree of stability across time and audiences such that the intervention is likely to remain effective. Thus, scaling is best used when there is a high degree of environmental stability. Of course, a strong relationship with key stakeholders will certainly help with scaling but weak relationships will not necessarily preclude it.

		Environmental stability	
		LOW	HIGH
Relationship strength	LOW	Dissemination	Dissemination but also consider scaling if an intervention is feasible without a strong relationship.
	HIGH	Knowledge translation	Both scaling and knowledge translation

Figure 1: When to use scaling, dissemination, or knowledge translation

SCALING

Behavioural science has had its biggest impact in the public policy sphere, where an understanding of behaviour influences how public decisions are made (Hallsworth & Kirkman, 2020). Governments are increasingly taking into account that much of individuals' behaviour occurs outside of conscious awareness and is driven by cues in their environment. Governments can better maximise public value by creating policy that accounts for actual human behaviour - indeed, this is the promise of behavioural public policy (Ruggeri, 2018). There are now more than 200 government Behavioural Insights teams across the world (Bhanot & Linos, 2020; Feitsma, 2019), demonstrating the importance and value in considering behaviour for designing policy and services. However, too often behaviourally-informed policy and research does not move beyond the "piloting" stage.

Taking a scaling perspective is one important way to ensure that you maximise the impact from the investment that you have made in understanding and testing behaviour change interventions. Most of the practitioners we speak to say that scaling is important to them and their organisations. It is common for end-of-project meetings to turn towards “next steps”, which invariably means “scaling”. Yet often, these conversations are ad hoc, and a side thought to the main project that is only considered at the end - it’s no wonder that more than 80% of effective innovations are not adopted at scale! As we will go through in detail below, scaling cannot be an afterthought. Taking a scaling perspective requires setting the scaling wheels in motion early.

What does scaling involve?

Researchers and practitioners use the term “scaling” in many different ways. For example, the World Health Organisation distinguishes between “vertical” scaling (e.g., changing institutions and policies) and “horizontal” scaling (e.g., expanding to new geographic locations) ([WHO/Expandnet, 2010](#)). MaRS Innovation Labs, a Canadian start-up and innovation hub, describe scaling “up” (e.g., changing institutions and policies), “out” (e.g., increasing the number of people reached by the intervention), and “deep” (e.g., changing cultural values and beliefs), and have created an online book and guide to explore these processes in detail (MaRS, n.d.). In our interviews with behavioural science practitioners that we undertook as part of the development of a “scale up toolkit for behaviour change” (more on this shortly), we found that scaling was used in similarly diverse ways: referring to increasing the **reach** of an intervention so that more people would experience the intervention; adapting an existing intervention to a novel **audience**, or adapting an existing intervention to a novel **behaviour** (Figure 2).

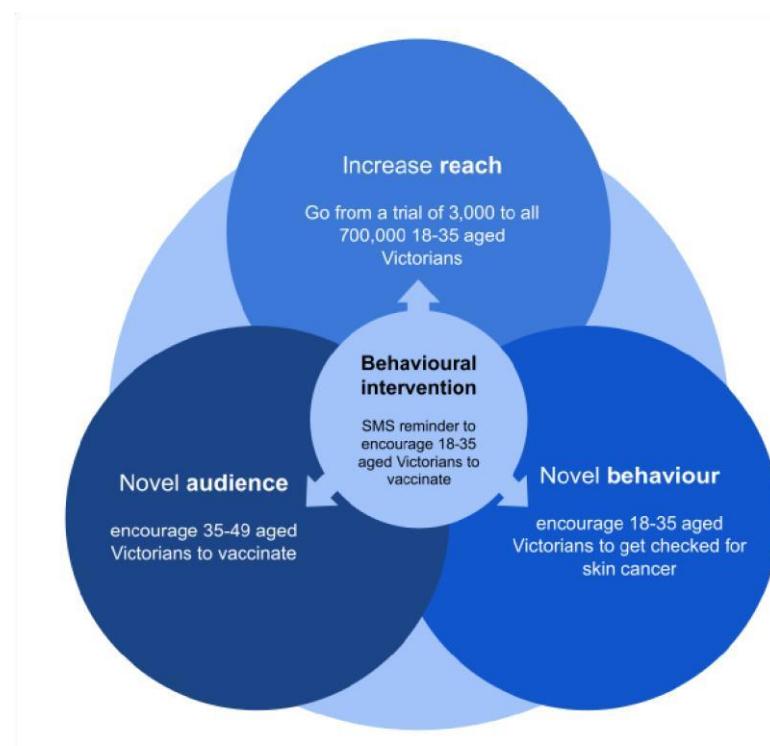


Figure 2: Perspectives on scaling from behaviour science practitioners

For clarity, our working definition of scaling up is “deliberate efforts to increase the impact of innovations successfully tested in pilot or experimental projects to benefit more people” (Saeri et al, 2021). This definition also draws on *Innovation and Scaling for Impact* by Seelos and Mair (2017), who describe scaling as “activities that act on and improve already existing knowledge, processes, products, services, or interventions to serve more people better” (Seelos and Mair, 2017, p. 7).

The two key elements in our definition of scaling are (1) the focus on scaling as an active process, and (2) the link between scaling and innovation, where innovation is the process of selecting or designing a novel behaviour change intervention to test in a pilot or trial and has the *potential* for impact (e.g., through estimating an effect size), and scaling is actually achieving that impact.

An intervention that is effective in a trial cannot be assumed to be effective when it is rolled out at scale, perhaps to a different cohort or a different context. One trap that researchers and practitioners can fall into is to emphasise novelty and innovation in a new intervention, product, or service, at the expense of consideration of its reach and impact (Al-Ubaydli et al., 2021).

Scaling challenges

Interventions may lack **external validity, so are limited to their testing or piloting context**. External validity is how well the results of a trial can be applied to other settings. This may refer to audience characteristics (e.g., it may work for older men, but not younger women), the organisational, regulatory or policy environment, or characteristics of the intervention itself.

Case study: Context matters for scaling prejudice reduction interventions

In a study that aimed to understand the characteristics of successful scaling of prejudice reduction interventions, it was found that context is critical (Hsieh et al, 2021). Across the 16 experts in developing and implementing prejudice reduction interventions, it was strongly emphasised that to successfully scale a prejudice reduction intervention, there is a need for genuine consideration of context. Contextual factors can include social dynamics, cultural, and linguistic factors. Similarly, Supplee and Kane (2021) highlighted that contextual factors like leadership, relationships with stakeholders and researchers, timing with political and policy windows - the realities of the policy process - need to be addressed. Indeed, as we consulted experts across the government, not-for-profit, private and academic sectors, our results go further to suggest that these are relevant across sectors and not just in the policy context.

Trying to scale up interventions can lead to a **loss of delivery fidelity**. Behavioural interventions sometimes rely on hidden idiosyncrasies of the testing context to work. Sometimes the importance of the context is not well understood because the interventions

that show promise in piloting also sometimes survive the replication test (trying the intervention again to see if it still works) because the replication context is highly controlled. But when the intervention is scaled and rolled out into the real world, those contexts aren't tightly controlled and interventions may flounder. Furthermore, it may be difficult for original members of the research team to be involved at roll-out sites to ensure critical delivery mechanisms can be repeated in other locations.

Implementation might need to occur nationally across many sites, so it is simply not feasible for the original team to be materially involved. Scaling up can sometimes be about handing over an effective intervention to users and organisations with expertise in service delivery and program management rather than behavioural methods. Balancing adaptability with rigour is a crucial challenge to ensure that an intervention's core "active ingredient" is maintained even when delivered by non-behavioural science experts.

A promising intervention may not be cost-effective or even effective when adapted to delivery at scale. That is, there could be **diminishing returns** to scale. This may be due to greater costs associated with training, the monitoring or evaluation of the people and organisations who deliver it at scale, or the costs and resources required to deliver the intervention to the most intractable segment of a population are disproportionately higher.

Finally, it can be particularly difficult to **mobilise and sustain commitment and resources** at scale. Mobilising commitment and sustaining commitment is about securing approval from decision-makers who would be responsible for implementing the intervention at scale. This is also affected by timing (e.g., government budget cycles), the political or organisational environment, or simply turnover in key relationships within partner organisations. The larger the scale and total budget required, the greater the scrutiny and perhaps challenges with securing resources.

These challenges are all intertwined: it can be difficult to generate and maintain commitment if an innovation is not cost-effective at scale, or appears irrelevant or a mismatch for the scaled-up context.

How to improve scaling?

It is important to start earlier than you might think when it comes to improving the scaling of behaviour change interventions. As several of the factors for successful scaling are related to the design of a behaviour change intervention, it is important to consider scaling as early as selecting a target behaviour (Chapter 5), and certainly by the time you select and/or design a behaviour change intervention (Chapter 8).

BWA and the Behavioural Insights Unit at the Victorian Department of Premier and Cabinet collaborated on a research project to answer the question: "How can we best support (behavioural insights practitioners in) scaling up?". We conducted an evidence and practice review to identify factors and activities that influence scale-up of behaviour change interventions, and developed and validated a free scale-up toolkit.

The BWA Scale-Up Toolkit includes five tools to help in the early stages of scale-up, each of which can be integrated throughout The BWA Method including:

- Identifying which behaviour to target with an intervention (Chapter 5)

- Assessing the feasibility of different intervention ideas (Chapter 8)
- Selecting a scalable behaviour change intervention (Chapter 8)
- Designing or adapting an intervention for testing and scaling (Chapter 11)
- Testing scaling assumptions about an intervention in a pilot or trial (Chapter 11)

You can access videos and tutorials on how to use the toolkit, and access extra resources and evidenceinformed guides to improve scaling at www.scaleup toolkit.org.

Case study: Scalability assessment for circular economy project

We conducted an evidence-based workshop series (“Circular Strategies”) to support organisational behaviour change and collaboration in the Australian textile, clothing and footwear ecosystem. This was one of three projects within the [BWA Waste and Circular Economy Collaboration](#), which was convened to look at the issue of waste and how to encourage Australians to avoid, reduce, reuse and recycle waste and/or adopt Circular Economy strategies.

We assessed the scalability of the workshop series for increasing Circular Economy practices in Australian businesses. While the project activities were focused on the Textile, Clothing, and Footwear sector, the ambition of the overall Collaboration was for the workshop series to act as a model for other sectors to adopt and adapt at scale.

We reflected on the activities we already conducted using the [scalability assessment tool](#) from the BWA ScaleUp toolkit. This is different to how the tool was originally envisioned, in which you would assess several candidate interventions and choose the most scalable. But using the tool in this way was still helpful for making assumptions about the path to scale up explicitly. It identified positives for scalability in that the workshops were delivered online and included several highly engaging activities that could be adapted to be standalone, and negatives for scalability in that participation in the program was maintained through significant (and costly) effort by the convenors and highly committed attendees.

The use of the tool helped to provide more useful recommendations to our government partners about how to encourage Circular Economy practices in Australian businesses, because it identified specific areas of benefit and concern for maximising the reach and impact of the workshop series as an intervention.

In addition to the intervention design and planning, it is often necessary to secure a commitment to scaling from a partner organisation as early as possible in the project, and maintain this commitment during the critical transition from trial to implementation. This helps to ensure that the necessary resources are available for scale-up when the time comes.

Finally, BWA has developed a tool to help you assess the scalability of an intervention and the capabilities of the teams and organisations that design, test, and scale up the intervention. We have summarised them according to the acronym SCALE (Table 1).

Table 1: SCALE factors relevant to scaling

Factor	Description
Simple	Scale-up is more effective when the intervention is <i>simple</i> and does not require specialised expertise, setting, or equipment for design, testing, or implementation.
Consistent	Scale-up is more effective when the intervention is tested in contexts <i>consistent</i> with those that it will be scaled into.
Adaptable	Scale-up is easier when the intervention is designed to be <i>adapted</i> : the core behaviour change mechanism is clearly specified, and adaptations, extensions and integrations with other interventions, programs and existing services are possible.
Likeable	Scale-up is more effective when the intervention to be scaled is <i>likeable</i> to the supporting organisations and target audience, such that they understand and endorse the intervention, and organisations specifically can integrate intervention delivery into their existing processes.
Effective	Scale-up is more effective when the intervention is perceived to be <i>effective</i> - and cost-effective - by those required to propagate and adopt it.

DISSEMINATION

Research dissemination is important because research will only influence behaviour if it finds its way to the relevant decision-makers and actors. Unfortunately, good research dissemination is rare. Practitioners often fail to disseminate their own research findings beyond their organisation due to having insufficient time, knowledge and incentives. In contrast, academic researchers are incentivised to disseminate their research findings via academic journals, which many practitioners report rarely reading or finding useful (Barraclough, 2004).

The rarity of good dissemination has significant social consequences because it slows the spread and implementation of important information into relevant settings. Key decisionmakers (e.g., policymakers, program managers) may make worse decisions because they lack relevant information. Those affected by their decisions (e.g., citizens, employees) may therefore suffer negative consequences that could easily have been avoided.

What does dissemination involve?

The aim of dissemination is to spread awareness of research insights and information. Dissemination activities include (1) strategising, planning and writing content for key outputs (e.g., publications, presentations, talks, policy briefs and social media) about research, and (2) communicating and sharing these outputs with key audiences (e.g., media, research users) who will (hopefully) consume and share them.

Successful dissemination usually requires tailored and strategic communication rather than simple distribution. For example, what resonates most with a team of policymakers might be very different from the ideal message to send to students. Dissemination to an academic

audience might require a credible academic source (e.g., a leading researcher), a comprehensive message (e.g., a paper), and a familiar channel (e.g., an academic journal). In contrast, dissemination to practitioners could require a very different set of sources (e.g., a respected organisation or domain influencer), message (e.g., a policy brief) and channels (e.g., email, social media and/or appropriate internal communication networks).

As the above suggests, successful dissemination involves effectively tailoring a range of variables rather than repeating a one size fits all approach for all audiences. The optimum strategy is the one that best matches all relevant variables. One approach for planning dissemination suggests a focus on four “components”: the source, message, channel, and audience (Figure 3).

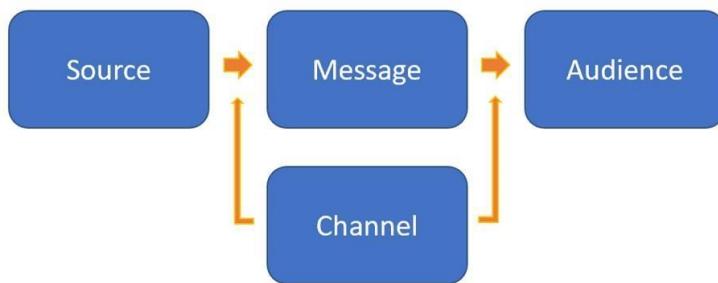


Figure 3: Key components in dissemination (adapted from Brownson et al., 2018).

The “source” is the recognised sender of the message. This might depend on age, gender, sex, ethnicity, credibility or trustworthiness. The “message” is the information (both verbal and non-verbal), which the source distributes. Such a message might differ in length, complexity and comprehensiveness. The “channel” is the means of distribution. This could differ in modality, credibility and reach (e.g., compare the implication of publishing a post on a journal, twitter or a popular blog). Finally, the “audience” refers to the recipients of the message and might, for instance, vary in engagement, knowledge and intelligence.

How to plan and execute dissemination?

Planning is essential for effective dissemination. There are a range of different frameworks that can be followed to plan research dissemination. These include [RE-AIM](#), the [Agency for Healthcare Research and Quality](#) and the Patient-Centered Outcomes Research Institute’s (PCORI) [Dissemination and Implementation Framework and Toolkit](#).

Research suggests several accepted principles to execute dissemination (e.g., RossHellauer et al., 2020; Ashcraft et al., 2020). These include general recommendations such as proper planning, resourcing and evaluation, but also picking insights specific to dissemination, such as the importance of adapting key messages and collaborating with other organisations and influencers.

To help you plan and execute dissemination, we have integrated the insights above into a single seven-point checklist (Table 2). For simplicity, we have ordered them under the abbreviation “SPREADS”.

Table 2: The SPREADS checklist for effective dissemination

Item	Explanation
Specification	I have selected the sources, messages, channels, audiences and outcomes that I will use
Preparation	I have found and engaged “amplifiers”: relevant individuals, organisations or networks who are, or can be, motivated to help disseminate this information.
Resourcing	I have evaluated resource needs and obtained appropriate skills and finances to support dissemination efforts.
Evaluation	I have determined when and how I will evaluate dissemination.
Adaption	I know how I will adapt my messaging as needed (e.g., for each amplifier, channel and audience).
Deployment	I have planned key actions and timeframes in response to the other prompts.
Sustainability	I have determined how I will maintain awareness and usage of the insights I disseminate.

Finally, in Table 3, we demonstrate how we used the SPREADS checklist when disseminating our Behaviour Change Scale-Up Toolkit, which we discussed earlier under “Scaling”.

Table 3: Applying the SPREADS checklist for effective dissemination to the BWA Scale-up project

Item	Explanation
Specification	We identified the most relevant academic and practical insights from our review, determined that we would target behavioural science practitioners, policymakers and academics working on scaleup, and do this using two different sets of channels.
Preparation	We found and engaged several individuals, organisations and networks who were interested in scale-up research and theory, and leveraged them to increase the reach and impact of our findings.
Resourcing	We evaluated resource needs and allocated resources for important aspects of our dissemination, such as website development and monitoring and writing an academic paper.
Evaluation	We informally evaluated the cost-effectiveness and impact of our dissemination strategies based on metrics such as visits and sign-ups.
Adaption	For practitioners, we developed a freely available Behaviour Change Scale-Up Toolkit. For academics, we wrote an article that is linked in the supporting materials of the toolkit. We produced a range of online and social media materials and delivered several presentations. We adapted a range of online and social media materials, and chose different sources and messages to optimise for different sub-audiences.
Deployment	We implemented and managed our dissemination plan and timeframes by using project management software (Asana).
Sustainability	We tried to sustain the impact of the toolkit in several ways, including (1) engaging practitioners over an extended period, (2) promising to develop future editions of the toolkit, and (3) providing updates about progress and related research.

KNOWLEDGE TRANSLATION AND BROKERING

Practitioners and public decision-makers work in complex, quickly changing environments, where they must draw on two kinds of “shortcuts” to help navigate the complexity of the real world. These can be “rational” shortcuts - pursuing clear goals and prioritising certain kinds and sources of information. Or in some cases, “irrational” ones - drawing on emotions, gut feelings, deeply held beliefs, favoured advisors, and habits (Burgman, 2016; Cairney, 2016; Cherney et al., 2015). Across public decision making, there can be a strong reliance on readily available, internal advice and personal networks to provide evidence and advice (Newman et al., 2016). Substantial challenges can exist connecting research and evidence to practical settings where it is able to have impact and vice versa. Consequently, there is a well-documented gap between the aspiration and reality of evidence-based public policy and administration (Cherney et al., 2015; Easton, 2018; Saltelli and Giampietro, 2017).

Relying only on the readily available evidence means that much of what is already known does not necessarily impact real-world practice. Often in public policy and administration, “the gap between what we know and don’t use is greater than the gap between what we know and don’t know” (Pupazzoni, 2020). It is understandable that decision-makers might take the path of least resistance by relying on the readily available, accessible, and digestible information. As social science entrepreneur Spencer Greenberg memorably commented when asked “doesn’t a focus on rapidly answering behavioural questions with large data sets and RCTs [randomised controlled trials] distort the kinds of interventions that people test or the kinds of questions that people ask in favor of things that are very cheap to do with a large sample?”, he replied: “we are like someone looking for lost keys in a dark room in many cases, so shouldn’t we at least start by looking under the lamplight before stumbling around in the dark?” (Wiblin and Greenberg, 2017).

To truly make the most of the available evidence, such as an effective intervention, there is a need to not just plan for its rollout at scale or to disseminate information about the intervention, but also to translate that information into available, accessible, and digestible forms, as well as deliberate effort to systematically embed knowledge in policy and program development. This practice is known as “knowledge translation” (Grimshaw et al., 2012) and “knowledge brokering” (Feitsma, 2019).

What does knowledge translation and brokering involve?

Similar to the previous point in our discussion of dissemination, for evidence to be influential, it is critical that it is available via trusted intermediaries (who may not always be scientists or academics), in relevant and accessible forms, at the right time, to meet the needs of decision-makers. This is where knowledge translation and knowledge brokering can be of assistance.

Knowledge translation has been increasingly used over the past 20 years to describe processes that ensure stakeholders are aware of and use research evidence to inform practice (Grimshaw et al., 2012). The term has been used interchangeably with several other terms (e.g., knowledge transfer and exchange, innovation diffusion, research utilisation, and evidence-informed policy - McKibbon et al., 2010). Knowledge translation offers an opportunity for organisations (e.g., government) to investigate whether there are insights or applied solutions that already exist, even if they come from other contexts. But it goes further too. Bammer (2013, p. 108) suggests four key translation goals:

1. Sharing what is known. This includes what has worked and has not worked, for whom, and why, so that policymakers and practitioners can develop effective actions.
2. Identifying remaining unknowns. This helps policymakers and practitioners take the unknowns into account, as well as to reduce, or at least be better prepared for, unintended consequences.
3. Critiquing current and proposed policy and practice.
4. Providing new ideas for policy and practice.

Bammer's full list requires three-way circulation between knowledge, policy, and practice, suggesting that successful translation needs to go beyond awareness and use of evidence, to also include relationships and interactions between major stakeholders, that is, knowledge brokering.

Knowledge brokering includes relationship development, ongoing communication and support, and often customised strategies that provide opportunities for individual and organisational capacity development (Dobbins et al., 2009). Knowledge brokers are just one term used to describe people who specialise in this type of activity. They are often, but not always, people in bridging roles and relationships between government and research institutions, whether they are part of an organisation and have strong connections with specialists like BWA, or BWA research fellows who embed into a partnering organisation (more on this later). They can act as a bridge between researchers and practitioners. Knowledge brokers also have a role to play not only in knowledge translation, but also in scaling and dissemination, where they can act as conduits for best-practice scaling and dissemination approaches.

How to improve knowledge translation and brokering?

The BWA Consortium and other long-term partnerships that we have entered into with government organisations are examples of ongoing knowledge translation and brokering. For many of BWA's Consortium partners, the preferred model has been "researchers in residence" where BWA staff would regularly attend and work in the organisation over months and years - being available for ad hoc advice and "hallway" conversations. This strategy can be useful particularly in organisations that do not yet have internal expertise on research in behavioural sciences. Over time, this can transition to mentoring and supporting internal champions and in-house experts, which also highlights the importance of capability building and learning. Under this model, BWA researchers can not only form a strong, on-the-ground understanding of the policy context and issues, but also develop an understanding of the organisational context and where and how to bring in research practices and evidence to strengthen the policy and decision-making processes and outcomes. Ongoing and long-term relationships between researchers, knowledge brokers and organisational stakeholders allow intractable problems and recurring challenges to be tackled in a much more iterative and exploratory way.

Case Study: Knowledge translation at The Shannon Company

Since 2011, BWA has partnered with The Shannon Company (TSC) - a leading communications agency and an original Consortium partner of BWA - to apply

theoryinformed research to influence policy and commercial campaigns. There is a dedicated BWA Partnership Manager, who supports the process of knowledge translation, and is regularly consulted on where behavioural insights and the academic literature can inform campaigns/communications. Over the past 10 years, TSC identified the benefit of upskilling their own team in behavioural science, which has led to the identification of bespoke training, and multiple members of the TSC team completing the BWA micro-credential course. The strong partnership between BWA and TSC has enabled the team to uncover valuable behavioural insights, particularly on emotional drivers, which has informed effective television and digital communication interventions for the likes of WorkSafe Victoria, the Transport Accident Commission, VicHealth, Industry Super Funds and Respect Victoria.

Of course, knowledge brokers do not have to be (BWA) researchers. With some of our other partnerships, we have worked with knowledge brokers who are staff of the partnering organisations, supporting them with specialist advice on particular behavioural science topics or running ad hoc “hackathons” to inform thinking on a policy issue with the pooled expertise of BWA.

Case study: Knowledge brokering at Sustainability Victoria

Sustainability Victoria was another original member of the BWA Consortium in 2011. It now includes several behavioural science specialists and a network of internal champions among their staff who act as knowledge brokers/translators, making the connection between research and practical application. This emerged from a 10-year portfolio of BWA research projects and training for Sustainability Victoria that has resulted in the behavioural sciences being embedded into organisational program design and delivery, as well as providing credibility when presenting these ideas to senior executives and Ministerial staff within the Victorian government. Sustainability Victoria is now recognised within the Victorian government as an agency that uses behavioural science to drive social change and impacts.

Knowledge brokering and translation is supported by shared understanding and shared capabilities. The BWA Method itself emerged from hundreds of projects with government partners, and a strong commitment, right from the start, to not just conduct important applied research, but also to build capability in behaviour change with government partners. By necessity, developing The Method also involved BWA staff learning to better understand how behaviour change projects fit into partner’s needs and agendas, and how findings can be integrated into organisational decision-making, programs and policies. Not surprisingly, given the challenge we acknowledge above with ensuring ongoing translation between research and practice, early forms of The BWA Method emphasised the research part, and not why it is needed, nor how it is applied, which BWA’s partnering organisations and cousers helped integrate. Building a shared understanding between BWA and its partners

(not to mention between nearly 30 behavioural scientists with diverse backgrounds and interests) via The Method has been another key value of long-term collaboration, knowledge brokering, translation, scaling and dissemination. And nothing teaches like teaching. Beyond the core Consortium, The BWA Method has been applied extensively throughout our training and education programs, including half day intensives with numerous government agencies, two-day masterclasses with the Australian and New Zealand School of Government, intensive “learn by doing bootcamps”, and a range of accredited masters units and microcredentials. The current version of The Method represented in all the chapters so far is a work in progress and no doubt there is much room to grow, but reflects a relatively mature method that has been applied across our projects with many different partnering organisations.

CONCLUSION

Returning to our case study from the start of this chapter, over the period of four years, our initial understanding of the problem shifted from identifying better practice signage and water quality risk communication, to conducting exploratory research into beach users’ decisionmaking processes in planning and preparing for a trip to the beach. This was informed through ongoing knowledge brokering and translation processes where a BWA knowledge broker worked closely with EPA Victoria over a number of years to understand the problem and draw in behavioural insights to inform and design the solution. The continued partnership and a willingness to test how behavioural science could be used to address the problem led to an experimental trial testing a beach water quality notification service to people’s mobile phones. After the trial, the service became mainstream into EPA Victoria’s beach water quality communications (scaling “out” to reach more people). Importantly, over this time we identified weather, and not warnings, was the strongest predictor of behaviours leading to exposure to poor water quality on bad days, thus the need for increased effort and activity when the weather is fine, but the water is not. As such, the risk assessment framework that was developed is now used to guide the response to highrisk days, where EPA disseminates water quality messages, designed using our research, via the print media, television, radio and social media channels. These are supported by interviews and public statements made by trusted public health authorities (e.g., EPA and /or the Victorian Chief Health Officer). Such a widespread campaign played a substantial role in the scaling out of the risk assessment framework.

Through the deliberate consideration of how to scale, disseminate information and translate knowledge, you can better ensure that an effective intervention and research evidence can reach the people that it needs to. In this chapter, we outlined how these three strategies of scaling, dissemination and knowledge translation help to improve implementation and how to conduct activities under each strategy. This careful consideration of implementation will ensure that effective interventions and behavioural science evidence can have the greatest impact possible.

So, does this chapter mark the end of The Method Book? Not yet. We have one final piece that will be released shortly, and will conclude and reflect on the process of documenting the BWA Method through the release of twelve chapters in twelve months (13 chapters if you count the Introduction we wrote in December 2020). We also have plans for, dare we say, to scale, disseminate and translate The Method Book in

2022 and beyond. So, there is one final chapter to go.

REFERENCES / FURTHER READING

Al-Ubaydli, O., Lee, M., List, J., Mackevicius, C., Suskind, D. (2021). How can experiments play a greater role in public policy? Twelve proposals from an economic model of scaling. *Behavioural Public Policy*, 5(1), 2-49. <https://doi.org/10.1017/bpp.2020.17>.

Al-Ubaydli, O., List, J. A., LoRe, D., Suskind, D. (2017). Scaling for Economists: Lessons from the NonAdherence Problem in the Medical Literature. *The Journal of Economic Perspectives: A Journal of the American Economic Association*, 31(4), 125–144. <https://doi.org/10.1257/jep.31.4.125>.

Agency for Healthcare Research and Quality (2014). Dissemination Planning Tool: Exhibit A from Volume 4.

Accessed on 26 November 2021.

<https://www.ahrq.gov/patientsafety/resources/advances/vol4/planning.html>

Anderson., D. (2008). Guest post: Proven programs are the exception, not the rule. Givewell blog. Accessed on: 26 November 2021. <https://blog.givewell.org/2008/12/18/quest-post-proven-programs-are-the-exception-not-the-rule/>.

Antman, E. M., Lau, J., Kupelnick, B., Mosteller, F., Chalmers, T. C. (1992). A comparison of results of metaanalyses of randomised control trials and recommendations of clinical experts: Treatments for myocardial infarction. *Journal of American Medical Association*, 268(2), 240-248. PMID: 1535110.

Ashcraft, L.E., Quinn, D.A., Brownson, R.C. (2020). Strategies for effective dissemination of research to United States policymakers: a systematic review. *Implementation Science*, 15, 89. <https://doi.org/10.1186/s13012-02001046-3>.

Balas, E. A., Boren, S. A. (2000). Managing clinical knowledge for health care improvement. Yearbook of medical informatics, 9(01), 65-70. PMID: 27699347.

Bammer G. (2013). Disciplining interdisciplinarity: Integration and implementation sciences for researching complex real-world problems. ANU Press, Canberra, <https://doi.org/10.22459/DI.01.2013>.

Barraclough, K. (2004). Why doctors don't read research papers. *BMJ*, 329(7479), 1411. 10.1136/bmj.330.7485.256-a.

Bhanot, S. P., & Linos, E. (2020). Behavioral Public Administration: Past, Present, and Future. *Public Administration Review*, 80(1), 168–171. <https://doi.org/10.1111/puar.13129>

Brownson, R. C., Eyler, A. A. Harris, J. K. Moore, J. B., Tabak, R. G. (2018). Getting the Word Out: New strategies for Disseminating Public Health Science, *Journal of Public Health Management and Practice*, 24 (2), 102-111. <https://doi.org/10.1097/PHH.0000000000000673>.

Burgman, M.A. (2016). *Trusting judgements: How to get the best out of experts*. Cambridge University Press, Cambridge, United Kingdom.

Cairney, P. (2016). *The Politics of Evidence-Based Policy Making*. Palgrave Macmillan UK, London.

Cherney, A., Head, B., Povey, J., Ferguson, M., Boreham, P. (2015). Use of academic social research by public officials: exploring preferences and constraints that impact on research use. *Evidence & Policy: A Journal of Research, Debate and Practice*, 11, 169–188.
<https://doi.org/10.1332/174426514X14138926450067>

Dobbins, M., Robeson, P., Ciliska, D., Hanna, S., Cameron, R., O'Mara, L., DeCorby, K., Mercer, S. (2009). A description of a knowledge broker role implemented as part of a randomised controlled trial evaluating three knowledge translation strategies. *Implementation Science*, 4(1), 1-9. <https://doi.org/10.1186/1748-5908-4-23>.

Easton, S. (2018). Fortress gov: Can scholars, bureaucrats team up to breach ministerial walls? *The Mandarin*. Accessed on 20 September 2018.
<https://www.themandarin.com.au/97308-fortress-government-academics-andbureaucratsmust-collaborate-to-penetrate-ministerial-offices>.

Feitsma, J. (2019). Brokering behaviour change: The work of behavioural insights experts in government. *Policy & Politics*, 47(1), 37–56.
<https://doi.org/10.1332/030557318X15174915040678>.

Grimshaw, J.M., Eccles, M.P., Lavis, J.N., Hill, S.J., Squires, J.E. (2012). Knowledge translation of research findings. *Implementation Science*, 7, 50.
<https://doi.org/10.1186/17485908-7-50>.

Hallsworth, M., Kirkman, E. (2020). *Behavioral Insights*. MIT Press, Cambridge MA.

Hsieh, Wickes, R., & Faulkner, N. (2021). What matters for the scalability of prejudice reduction programs and interventions?. <https://doi.org/10.31234/osf.io/h6p8e>.

MaRS. (n.d.) Living Guide to Social Innovation Labs. Accessed on 26 November 2021.
<https://mars-solutionslab.gitbook.io/living-guide-to-social-innovation-labs/>.

McKibbon, K.A., Lokker, C., Wilczynski, N.L., Ciliska, D., Dobbins, M., Davis, D.A., Haynes, R.B., Straus, S.E. (2010). A cross-sectional study of the number and frequency of terms used to refer to knowledge translation in a body of health literature in 2006: A Tower of Babel?. *Implementation Science*, 5:16. doi: 10.1186/1748-5908-5-16.

Morris, Z.S., Wooding, S., Grant, J. (2011) The answer is 17 years, what is the question: understanding time lags in translational research. *Journal of Royal Society of Medicine*, 104(12), 510-20. doi:10.1258/jrsm.2011.110180 .

Newman, J., Cherney, A., Head, B. (2016). Do Policy Makers Use Academic Research? Reexamining the “Two Communities” Theory of Research Utilization. *Public Administration Review*, 76, 24–32. <https://doi.org/10.1111/puar.12464>

Patient-Centered Outcomes Research Institute (2016). Dissemination and Implementation Framework and Toolkit. Accessed on 26 November 2021.

[https://www.pcori.org/research/putting-evidence-work/disseminationandimplementationframework-and-toolkit.](https://www.pcori.org/research/putting-evidence-work/disseminationandimplementationframework-and-toolkit)

Pupazzoni, R. (2020). Fiona Stanley on how she became the driving force behind folate for pregnant women. ABC News. Accessed on 26 November 2021.

<https://www.abc.net.au/news/2020-05-23/fiona-stanley-womanbehind-folate-to-stop-spinabifida/12276118>

Re-Aim. Accessed on 18 November 2021. <https://re-aim.org/learn/what-is-re-aim/>

Ross-Hellauer, T., Tennant, J.P., Banelytē, V., Gorogh, E., Luzi, D., Kraker, P., et al. (2020) Ten simple rules for innovative dissemination of research. *PLoS Computational Biology* 16(4): e1007704. <https://doi.org/10.1371/journal.pcbi.1007704>.

Ruggeri, K. (2018). *Behavioral Insights for Public Policy: Concepts and Cases*. Routledge. <https://doi.org/10.4324/9781351052542>.

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

Seelos, C. & Mair, J. (2017). *Innovation and Scaling for Impact: How Effective Social Enterprises Do It*. Redwood City: Stanford University Press.
<https://doi.org/10.1515/9781503600997>

Saltelli, A., Giampietro, M., (2017). What is wrong with evidence based policy, and how can it be improved? *Futures*, 91, 62–71. <https://doi.org/10.1016/j.futures.2016.11.012>

Supplee, L., Kane, M. (2021). The realities of scaling within evidence-based policy. *Behavioural Public Policy*, 5(1), 90-102. doi:10.1017/bpp.2020.25

Westfall, J. M., Mold, J., Fagnan, L. (2007). Practice-Based Research - “Blue Highways” on the NIH Roadmap. *Journal of American Medical Association*, 297(4), 403-406. doi:10.1001/jama.297.4.403

Wiblin, R., Greenberg, S., (2017). Speeding up social science 10-fold, how to do research that's actually useful, & why plenty of startups cause harm. 80,000 hours. Accessed: 25 March 2021. <https://80000hours.org/podcast/episodes/spencer-greenberg-social-science/>.

World Health Organization/ExpandNet (2010). Nine steps for developing a scaling-up strategy. Accessed on 26 November 2021.
https://www.who.int/reproductivehealth/publications/strategic_approach/9789241500319/en/