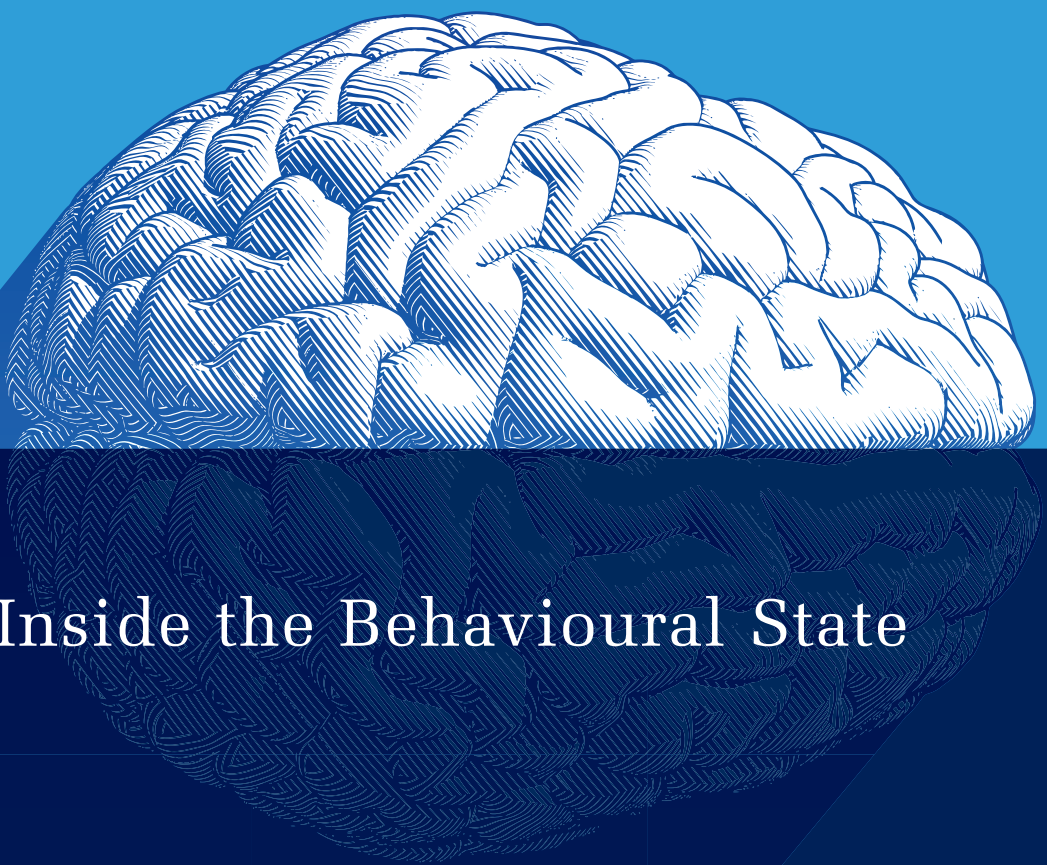


# Inside the Behavioural State

Joram Feitsma

eleven
international publishing

Inside the Behavioural State

**Promotiecommissie**
Prof. dr. A.J. Meijer
Prof. dr. M. Noordegraaf
Prof. dr. H. Strassheim
Prof. dr. W.A. Trommel
Prof. dr. B. Vis

# Inside the Behavioural State

Binnenin de Gedragsoverheid
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. H.R.B.M. Kummeling, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op donderdag 27 juni 2019 des middags te 6.00 uur.

door

Joram Nanne Pieter Feitsma

geboren op 21 september 1990
te Amsterdam

**Promotoren**
Prof. dr. T. Schillemans
Prof. dr. P. ’t Hart

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van NWO in de vorm van een TOP-subsidie voor het onderzoeksproject Welfare Improvement Through Nudging Knowledge (WINK).

# Acknowledgements

For this doctoral study I have studied behaviour experts in the public domain from up close. Going 'inside the behavioural state' has been a rich journey, which has taught me a lot about 'doing' science, about public policy and public administration, and most particularly: about the still fascinating world of public behaviour change. This book will introduce you into this world, full of ambition and opportunity but also ambiguity and tension. Before we begin however, I'd like to express my appreciation for the people whom I've had the great pleasure and luck to be working together with along the journey. I owe an especially big 'thank you' to:

Conversation and collaboration partners in the behavioural policy field. Thank you all for your willingness to share your stories. Giap, you introduced me into the behavioural scene at the ministries. I've enjoyed our conversations, fancy dinners and conference meet-ups. Jelle and Sophie, thanks for the pleasant collaboration. The 'Urban Nudging' project was a great starting point to learn about behaviour change beyond the ivory tower. Harm and Odette, thank you for providing me the opportunity to tag along with your team. It provided a vital ethnographic experience.

Fellows in the WINK research group and consortium. It has been great working together with you over the years. I look forward to celebrating our project with the upcoming WINK conference in June. Tina and Stas, 'WINK jr.', although we were based at different departments, having you as fellow PhDs made my own project feel less isolated. I wish you the best of luck in finishing up. Jorren, I've enjoyed working together. Niek, our lunch updates were always good fun. You being one of my chief 'insiders' in the field as well, it's an honour to have you spin the tunes at the defence party.

International academic colleagues. Mark, thank you for making me feel at home, also academically speaking, at Aberystwyth University during my research stay in the fall of 2017. I will not soon forget our inspiring chats and paper discussions, let alone that epic 'Twin Peaks' run. Sarah, in you I found a first fellow 'BITnographer'. A surprising and encouraging experience. I'm glad I made a visit to your university in Brisbane last summer. Our sparring and 'shut up and write' sessions made for a productive but beyond that also a refreshing collaboration.

Colleagues at Utrecht School of Governance, NIG comrades, Inquisitive PA pals, conference amigos, and office mates. You've made for a highly pleasant and stimulating social work environment. Fellow PhDs at USG, thank you for the energizing lunch walks, sparring sessions, paper discussions, coffee talks, *vrijmibo's*, long train rides, dinners, and of course most importantly: the annual laser tag events. The terminator is eager to see you on the battlefield again. Jo, Lauren, Marija, Erik-Jan, thank you for your helpful feedback and kind help with translating as well. A few people deserve special mention. Erik-Jan, *paranimf* and kindred spirit, having you around in the office has been a great source of joy. Lars, I've enjoyed our inspiring chats, hikes and dinners over the years. Peter, you sparked the academic flame in me, thank you. Hans, I've learned a lot from you about teaching and training over the years, your considerateness is inspiring. Martijn, you provided indispensable support at crucial times along the journey, thank you.

Those most close to me. Tom, Piter-Jan, Oskar, Leonie, Erik, thank you for your invigorating presence in my life. Wolves, let's howl on. Mila, thank you for having made the journey so much warmer and lighter. Jelmer, it's great to have you as *paranimf*, even better as inspiring brother. Jeanet and Theo, your loving support has always been there. It is to you that I dedicate this book.

Last but certainly not least, I am deeply grateful for my promoters. Paul, my 'thunder strike promoter', coming in with blasts of insights which although often confronting have helped advance this book so crucially. Thank you, Paul, for your critical thinking and help along the journey. Thomas, my academic mentor, from you I've learned so much it's hard to know where to begin. I wholeheartedly thank you for your abundant knowledge sharing and close involvement over the years.

Joram Feitsma,
Utrecht, May 2019

# Table of contents

# I | Introduction

*Chapter 1*

# Seeing like a Behavioural State

## 1.1 Initiation into the behavioural state

***The emerging behavioural state: promise, pressure or phantom?***

How confident is present-day government? Not too confident, it would seem. Solving public problems is generally seen as an arduous challenge. Our late modern society doesn't lend itself to be 'engineered'. It has only become more uncertain, complex, and 'liquid' (Bauman 1999) than before (e.g. Beck 1992; Giddens 1998). This acknowledgement of complexity reflects how today's policymakers think, 'see' and act. They commonly acknowledge that some policy issues have become 'wicked' (Head 2008): so convoluted and contested – both substantively and procedurally, empirically and normatively – that they can at best be 'tamed'. Also, today's policymakers – part of the *New Public Governance* paradigm – generally realize that they are heavily dependent on other societal actors to get things done and that forging strong and unified public-private networks is far from easy. So they buckle up, strengthen their tolerance for uncertainty and nonperformance, and confess that public problem-solving simply isn't easy.

And yet, in the past two decades a policy trend has emerged that goes in the opposite direction of a modest state. This trend has given birth to a type of state that seems all too confident about its ability to shape society. This confidence is drawn from the opportunities discovered by recent advances in the behavioural sciences. A body of 'behavioural insights' (e.g. Dolan et al. 2010; OECD 2017; Lourenço et al. 2016) has developed rapidly, showing how humans often don't make decisions in wholly rational, but often more automatic ways. These insights have given rise to an unorthodox policy style, which subtly intervenes in the everyday spaces of society based on rigorous behavioural analyses and experiments (e.g. Jones et al. 2013; Leggett 2014; Strassheim et al. 2015; Whitehead et al. 2017; John 2018). Moreover, to incorporate these insights, a new professional policy community has emerged onto the global policy scene. The rise of this community – what I coin 'the behavioural state' – is the central topic of this study.

This emerging behavioural state has potentially radical, though unclear, ramifications for the policy world and its impacts on society. Three radically different expert viewpoints prevail. On one hand, the behavioural state is viewed as a grand *promise* (e.g. Dolan et al. 2010; World Bank 2015; OECD 2017). Advocates of the behavioural state claim that new behavioural science insights into the bounded rationality of humans have the potential to increase the performance of governments. They foresee cheaper, more effective, evidence-based policy solutions. They foresee that even the most 'wicked' issues can be solved, from obesity to climate change. Moreover, advocates cherish the 'soft' nature of the behavioural state's intervention techniques, which typically don't enforce choices but merely alter the context in which these

choices are presented. They softly 'nudge' (Thaler & Sunstein 2008) people towards desired behaviours. These techniques are celebrated as an attractive middle ground between laissez faire liberalism on one side and heavy-handed state regulation on the other. In fact, advocates claim that they don't just respect individual freedom of choice, they actually empower citizens to make choices in their long-term interests. They 'help people help themselves' (Loewenstein et al. 2012).

At the same time, for some the behavioural state is a cause for serious concern. Dystopian accounts exist that portray the behavioural state as a form of 'late modern *pressure*' (Van Putten & Trommel 2018; emphasis JF) that threatens democratic procedures and the liberal freedoms of citizens. Critics warn against deepened forms of paternalism (Furedi 2011), increased government manipulation and the resurgence of technocracy (White 2013). They foresee the rise of powerful systems of 'psychopolitics' (Han 2017) and 'psychocracy' (Jones et al. 2013) that seek to regulate and discipline the public with subtle psychological behaviour change techniques. Other critics argue instead that the behavioural state *lacks* power, as it suffers from an obsession with psychology and a rationalist hubris (Lodge & Wegrich 2016).

In addition, the debate is coloured by deep scepticism, dismissing the behavioural state as a neither substantive nor innovative development (e.g. Ryan 2017). Sceptics view the widely declared 'behavioural turn' as a mere whim, a rhetorical *phantom* haunting the policy world with inflated concepts but little substance in practice. Lacking critical substance, this phantom can neither increase governments' performance nor threaten citizens' liberal freedoms. Sceptics seek to expose the hype and bring the aggrandized expectations about the behavioural state to an end. After all, there are real innovations that need to be addressed.

This contradictory set of viewpoints leaves us puzzled. How can the behavioural state be explained as promise, pressure and phantom all at once? Is it because these different camps in the debate – advocates, critics and sceptics – appraise the behavioural state from different disciplinary or normative perspectives? Undoubtedly they do, and this is only logical. However, a more problematic question is: are these camps actually talking about the same empirical phenomenon in the first place? I would beg to differ. My main claim is that the behavioural state has been overtheorized but empirically understudied. We have only seen a tip of the iceberg. My main proposition is that, before jumping to grand conclusions, a closer look is needed into the concrete shapes that the behavioural state has taken. We need to dive in deeply and ask: what/who/where is it, really, this behavioural state?

***Capturing the behavioural state***

The simplest and most conventional way to describe the behavioural state is to give illustrations of its associated programs and policies. Below are three examples.

> *Tobacco Display Ban.* In 2015, the UK government introduced a new regulation that prohibited small shops from displaying cigarette products to customers. Products had to be placed behind closed counters and could only be displayed temporarily on the point-of-sale. The idea was that making tobacco products less visible would denormalize smoking and prevent young adolescents from taking up smoking.
>
> *Healthy School Canteens.* Dutch secondary school canteens are nowadays inspected by the governance-financed Netherlands Nutrition Centre. 'School Canteen Brigade Officers' check whether these canteens steer schoolchildren towards healthy food choices. If needed, they suggest all sorts of 'nudges' (Thaler and Sunstein 2008), such as hanging posters of healthy food, presenting fruit more attractively, and placing healthy product options more within reach. Schools deemed healthy enough are publicly awarded a 'Healthy School Canteen Golden Plate'.
>
> *Expedition Work.* In 2012, the Dutch Employee Insurance Agency launched 'Expedition Work', a serious gaming environment that helps jobseekers become more employable. 'Expedition Work' consists of an online dashboard with various games in which jobseekers can do self-tests and practice skills related to jobseeking. The more games they play, the more points they earn, which are displayed in their personal score.

These illustrations give a quick sense of the behavioural state. To begin with, they reflect how the behavioural state 'sees'. This state has an acute awareness of the *role of the human unconscious*, acknowledging that humans are not single-minded rational agents but instead easily influenceable, fallible decision-makers who don't always act in accordance with their long-term interests (Kahneman 2011; Dolan et al. 2010). Most of all, the illustrations demonstrate how the behavioural state acts. Its way of intervening reflects a *softness*, not obliging but merely 'seducing' people into desired behaviours. This seduction works by subtly readjusting people's daily environments, also known as 'choice architecture' and 'nudging' (Thaler and Sunstein 2008).

And yet, there is much more to the behavioural state than its nudges alone. Consider, for instance, the academic research available. Thus far, the behavioural state's emergence has spurred diverse lines of scientific inquiry. Behavioural scientists have studied the working mechanisms and effectiveness of behaviour change interventions in numerous policy domains (*inter alia,* Arno & Thomas 2016; Bucher et al. 2016). Political philosophers and theorists have discussed the ethics and politics of the

behavioural state (*inter alia,* Bovens 2009; Leggett 2014; White 2013). Scholars of law have considered the implications of behavioural public policy for the judicial system (*inter alia,* Allemano & Sibony 2015; Van Aaken 2015; Lepenies & Malecka 2015). Social and political geographers have performed 'nudgeography' (Reid & Ellsworth-Krebs 2019), looking at how the behavioural policy agenda has been spreading across the global policy sphere (*inter alia,* Jones et al. 2013; Jupp et al. 2016; Whitehead et al. 2014). Public policy scholars have explored how behavioural policymaking is being institutionalized, mostly focusing on activities in the Anglosphere (*inter alia,* Strassheim & Korinek 2014; John 2018).

Realizing that these different disciplinary lines of inquiry can offer at most partial insights, more recently also multi/cross/interdisciplinary approaches have emerged. Academic journals have been launched with an explicit interdisciplinary outlook, such as Behavioural Public Administration and Behavioural Public Policy - which goes as far as welcoming 'economists, psychologists, philosophers, anthropologists, sociologists, political scientists, primatologists, evolutionary biologists, legal scholars and others' (Behavioural Public Policy Website 2018) to submit articles. Additionally, interdisciplinary research networks have been formed, such as the Interdisciplinary Behavioural Insights Research Centre at Aberystwyth University, or the Behaviour and Institutions Stream at Utrecht University. And lastly, various individual research programs have adopted a multi/cross/interdisciplinary perspective. This study is the product of such a program, named Welfare Improvement Through Nudging Knowledge (WINK).[1]

Notwithstanding the importance of these diverse and increasingly coupled research perspectives so as to transcend a narrow understanding of the behavioural state, I contend that the current research agenda suffers from a more urgent issue: the lack of empirical and practice-oriented research approaches. To see this, let's return to the puzzling mix of conflicting appraisals of the behavioural state. Advocates view this state as a great promise for a more effective and efficient liberal government. Critics on the other hand see it as either powerless due to its alleged methodological selectiveness, or too powerful, featuring technocratic practices that manage to subtly regulate human conduct at the cost of democracy and autonomy. Sceptics, on another side, dismiss it as an inflated, fuzzy or marginal phenomenon. Which of these

[1] WINK's goal is to explore the merits of behavioural public policy as an innovative approach to public policymaking, with a focus on public health and welfare (see WINK Website 2019). Research is undertaken from three disciplinary perspectives: social health psychology, political philosophy and public administration. This particular study takes the public administration perspective. WINK's overarching aim is to *connect* questions related to the behavioural state's effectiveness, ethics and feasibility that are otherwise only studied in isolation.

radically different views are accurate? As things stand, I would contend that we can't properly tell. That is because these views seem to reside from a debate that misses a strong empirical foundation. Studies have tended to be abstract, discussing the *idea* and *ideal* of behavioural policymaking, typically focusing on interventions as such. The debate lacks deep empirical knowledge of the behavioural state. Particularly, it lacks knowledge of its many members and their daily practices (Lepenies & Malecka 2018), who together form a so called 'Behavioural Insights' community. Much about these people – the self-proclaimed 'Behavioural Insights Teams' or 'BITs', 'Nudge Experts' and 'Choice Architects' – is unknown. Who are they? What does an average day for them look like? Which ideas inspire them? What drives them? Who are their heroes? How do they 'see'? What kind of techniques do they use? Do they follow particular routines and procedures? Rituals too perhaps? What is their typical jargon? Such questions are rarely asked.

This 'anthropological blind spot' within the behavioural state debate is unfortunate, especially as learning more about this new occupational group in government might well help us to figure out the puzzle provided by the deadlocked debate between advocates, critics and sceptics. To explore what this behavioural state actually entails – promise, pressure or phantom – this study therefore aims to remove this blind spot. It will shift the focus to the behavioural state's 'early adopters' in order to learn from them about what this state really has to offer (or not). It will present an in-depth account of actual behavioural state practices, focusing particularly (but not exclusively) on practices within Dutch government. How this study will do so, and by which questions it will be guided, is explained in the next section.

### 1.2 Method matters

***Main research aims***

This study seeks to capture the behavioural state empirically, both in its richness and realness, looking further than solely its public impressions and emblematic policies. The behavioural state will be studied from an anthropological angle – as a still evolving professional community with its own ideas, methods, routines, techniques and symbols. The aim is to learn from this Behavioural Insights (hereafter: BI) community about behavioural state *practice*. By following this BI community in its frontstage- but also its backstage developments (Goffman 1959), we can gain insights into how established the behavioural state actually is, where it is located, by whom it is inhabited, how these people 'see' and what they actually do.

The goals of this study are both descriptive and evaluative. It starts out with a more descriptive ambition, seeking to uncover actual behavioural state practices. From there, the study will also discuss how this all *matters* for state and society. It seeks to identify the real merits and risks of the behavioural state. 'Big' questions like 'Where

is this going?', 'Is this a desirable development?', and 'Could it be different?' (Flyvbjerg 2001) will be explored.

***The need for empirical depth and breadth***

As hinted at already, the behavioural state should be studied more closely because the current behavioural state debate seems confused. Although the BI movement has a seemingly simple and promising message, it has still elicited quite a polarized debate (Towfigh & Traxler 2016). It has become the subject of an academic 'trench warfare' in which it is simultaneously appraised as a promise, pressure and phantom. This puzzling mix of viewpoints at the least suggests that this behavioural state has a rather multifaceted nature, and that behind the gloss of the 'golden concept' (Pressman & Wildavsky 1974) lies a host of ambiguities and tensions.

One may question whether all of these ambiguities and tensions are readily acknowledged in the behavioural state debate. My contention would be that the debate is insufficiently informed. To begin with, it would benefit from more empirical depth. Advocates have tended to reiterate a grand narrative about 'the promise of Nudge'. They litter their writings with promising examples of subtle nudges with strong effects. Prototypical are the update reports in which the interventions and outcomes are shared by forerunners in the BI field (e.g. Van Bavel et al. 2013; Lourenço et al. 2016; OECD 2017). A narrow focus exists on interventions, and in fact, a limited selection of emblematic success stories. Critics too have tended to focus on interventions, while also making big and abstract leaps from small nudges to fundamental theories about autonomy, democracy and the legitimacy of state intervention (e.g. Button 2018; Hansen & Jespersen 2013). Meanwhile, sceptics' radical claims about BI's novelty and impact have also tended to lack a thorough empirical basis that includes knowledge of current developments in the field (e.g. Fitzpatrick 2011; Selinger & Whyte 2012). The result is that we have mostly only seen the behavioural state's end output and outcomes. We now know what kinds of programs and policies it ultimately produces, but we remain ignorant about how these actually come about. In other words, we have seen the policies but not the policy*makers* and the policy*making*.

Besides depth, the debate also needs more empirical breadth. Thus far, it is mostly the high-profile, trendsetting Anglo-Saxon units – the so called 'Behavioural Insights Teams' or 'BITs' – that are being mentioned and investigated (e.g. John 2014; 2015; 2018; Ball et al. 2017). A bit of academic research has looked into life inside BIT UK, the primary trendsetter in the field, with by now over hundred employees and offices in the South East Asia, Australia and North America (John 2014; 2015). Developments within the Behavioural Economics Team of the Australian Government (BETA) have been analysed as well (Ball et al. 2017). Still, most of what we know about these Anglo-

Saxon units comes from non-academic publications from these units themselves (e.g. Dolan et al. 2010; Haynes et al. 2012; Service et al. 2014; Hallsworth et al. 2018; Halpern 2015). Beyond this Anglo-Saxon geographical foothold, the BI landscape is relatively opaque. Little research has been undertaken into, for instance, the Qatar Behavioural Insights Unit, Nudge Unit Greece or Behavioural Insights Group Rotterdam.

This study intends to bolster the debate's currently fragile knowledge base with more empirical knowledge of behavioural state practice. Firstly, empirical breadth will be provided by shedding light on a site that has thus far been little explored: the BI landscape within Dutch government. Studying the Dutch case will help to understand how BI is institutionalizing beyond the Anglosphere, and potentially see how aspects of local institutional context may influence behavioural state practice. Moreover, the Dutch case provides us with a rich landscape of exploratory practices, both at the central and local level. This varied and still evolving landscape offers a great starting point for in-depth empirical research. Secondly, empirical depth will be provided by shifting the focus away from the policies made: the well-known, emblematic success-stories presented at the frontstage. Instead, this study zooms in on the makers and the 'making of' of those policies, i.e. on the BI movement's backstage activities. The specific research foci are clarified below.

***Research themes and questions***

As indicated, this study seeks to uncover actual behavioural state practices. To do so, it follows and seeks to learn from ongoing activities in the field. It takes as its main study object the self-professed experts that see themselves as heralds of the BI trend. The broad research question is:

> *What can we learn about the behavioural state by studying its members and their practices from up close?*

This question can be broken down into three lines of empirical inquiry. A first line of inquiry looks into *practices* of the behavioural state, and even more basically, its prevalence. Where can we find behaviour experts in government? Who are these people and what are their professional backgrounds? How are they organized? What is being done and accomplished? Investigating these basic questions is important if we take seriously the sceptical concerns about the marginal or even merely rhetorical nature of BI. So, as an overarching sensitizing question, we ask: *is there a behavioural state?*

A second line of inquiry studies how behaviour experts are developing collectively. The core theme here is that of *professionalization,* dealing with how the BI community is establishing its own standards and how it is fitting in within the established policy order. This theme is important given that the future of BI is still open-ended and its

impact uncertain. From the outset, BI still seems relatively marginal, and also reliant on a *Nudge* trend that might eventually prove unsustainable. Also, it seems relatively unrestricted by internally or externally imposed institutional standards. A question is whether more strict professional boundaries will be established in the near future, and if so, what they will entail. Against this background, BI's developmental trajectory needs to be studied more closely. How is this new area of expertise being institutionalized? Is this being done successfully? What is specific about Dutch behavioural expertise – in comparison with international practice? In other words, we ask: *how is the behavioural state developing?*

A third line of inquiry concerns the role of *expertise* within the behavioural state. This theme is important given that the role of knowledge in public policymaking is far from clear-cut. Policies are typically developed against a background of uncertainty, and there are many – both overt and covert – forms of 'misuse' of knowledge (Cairney 2016). Such circumstances make the seemingly simple idea of introducing behavioural expertise in policymaking less obvious. It is thus relevant to capture how behaviour experts handle evidence in a context of uncertain, ambiguous and/or contested knowledge. What type of rationalities do these experts subscribe to? Which methods and theories do they favour, or discredit instead, and for what reasons? Put broadly, and borrowing from James C. Scott's (1998) terminology, we ask: *how does the behavioural state see?*

Together, the empirical chapters zoom in on actual practices (Chapters 3 and 4), professional developments (Chapters 5 and 6), and the role of expertise (Chapters 7, 8 and 9) within the Dutch behavioural state. Table 1.1 summarizes the research questions, associated topics and chapter contributions.

**Table 1.1: Three lines of empirical inquiry**

| *Broad question* | What can we learn about the behavioural state by studying its members and their practices from up close? | | |
|---|---|---|---|
| *Line of inquiry* | *Practices* | *Professionalization* | *Expertise* |
| *Sensitizing question* | Is there a behavioural state? | How is the behavioural state developing? | How does the behavioural state see? |
| *Topics* | Prevalence<br>Size, scope and whereabouts<br>Daily activities | Institutionalization<br>Uniformity<br>Boundary formation<br>Policy translation | Institutional logics<br>Deliberative qualities<br>Knowledge use<br>Interdisciplinarity |

| | | | |
|---|---|---|---|
| *Contribution of the empirical chapters* | - Chapter 3 explores BI practices at the central state level<br>- Chapter 4 explores BI practices at the local state level | - Chapter 5 analyses BI's professional developments at the central state level<br>- Chapter 6 analyses BI's professional developments in international context | - Chapter 7 analyses institutional logics underpinning BI<br>- Chapter 8 analyses alternative modes of expertise within BI<br>- Chapter 9 analyses varied degrees of interdisciplinarity in BI |

***'BITnography': the ethnographic study of the behavioural state***

To move beyond the empirically limited scholarly 'trench warfare' about the behavioural state, this study shifts the debate's focus from studying programs and policies to studying people and practices. Not nudg*es* will be studied but nudg*ers:* the self-professed champions of the current BI trend. By following their movements, we can come to see how BI is actually practiced, in all of its richness and realness.

My core method for this study has been doing *ethnographic fieldwork.* Ethnography, very roughly, is about coming to understand a particular group of people by spending time with them for a longer time within their natural habitats (see, *inter alia,* Agar 1996; Hammersley & Atkinson 1983; Rhodes et al. 2007; O'Reilly 2008). It is about observing and participating with them from the inside, as a route towards apprehending their group culture – whether that is the culture of homeless persons in San Francisco (Gowan 2010), sex workers in the favelas of Rio the Janeiro (Silva 2015), or ministers in British government (Rhodes 2011). By immersing oneself in an unfamiliar world, the ethnographer tries to go beyond the fabricated façade of people or organizations and find out what really occurs 'backstage' (Van Hulst 2008; Goffman 1959). Likewise, I have immersed myself in the behavioural state, seeking to understand it from the inside. I have gotten 'up-close and personal' (Rhodes et al. 2007) with a pioneering community of behaviour experts over the course of four years, from September 2014 to December 2018.

Broadly speaking, this study can be positioned within a body of ethnographic studies in the field of public administration (hereafter: PA). 'Administrative ethnography' (Boll & Rhodes 2015) is a relatively marginal but growing body of research (Cappellaro 2017). Recent important works promoting ethnography in the field of PA and political science come from Edward Schatz (2013), Dvora Yanow (2000) and Rod Rhodes (Rhodes et al. 2007). Implicitly, this type of research could also be associated with the work of some leading PA scholars, like Herbert Kaufman (1981) and Aaron

Wildavsky (Heclo & Wildavsky 1974), who also conducted immersive, long-term, observation-based research.

In a narrower sense, this study falls within an even more modest body of qualitative and ethnographic studies on the behavioural state: 'BITnography', so to speak. Key work within this body includes that of Peter John (2014), who has spent time with the pioneering BIT UK and used his observations to analyse the unit's developments and successes. More recently, John has also looked into the global spread of BI (John 2019). Additionally, Sarah Ball (Ball et al. 2017) has translated her participant-observations at the Australian BETA into practical lessons for setting up a BIT. Colette Einfeld (2018) has interviewed Australian policy workers involved in BI projects, asking about their views on how novel Nudge is and how it relates to evidence-based policymaking. Mark Whitehead and colleagues (2017) have traced the rise of what they call the *neuroliberal,* and before that the *psychological* (Jones et al. 2013) state, based on a rich document study and numerous expert interviews. This study seeks to contribute to and strengthen this body of research by exploring a thus far little explored site of BI (i.e. the Dutch BI landscape) based on relatively 'thick' data (i.e. studying, interviewing and observing Dutch behaviour experts over the course of four years).

***Critical-constructive fieldwork***

The fieldwork of this study has both critical and constructive elements in it. Its critical nature comes from its connections with the critical public administration literature (*inter alia,* Forester 1993; Box 2005; Flyvbjerg 2001; Frissen 2013; Trommel 2009; Ossewaarde 2010; Kruiter 2010). According to critical thinkers in the field of public administration, the problem of mainstream PA schools is that they have grown too much in a consultancy role, assisting existing power structures without fundamentally reflecting on the values and interests that they serve and strengthen (Van Putten & Trommel 2018). The types of knowledge that mainstream PA produces tend to be relevant for either a professional audience (expanding on current research and refining the methodology) or a policy audience (developing new tools and knowledge that can be directly used to manage issues). The knowledge produced is of an instrumental nature. According to critical PA scholars, mainstream schools have lost touch with the field's original adage of 'speaking truth to power', as formulated by Aaron Wildavsky (Van Putten & Trommel 2018). Critical PA distances itself from the mainstream schools by generating a more normative, 'critical-reflexive' (Schillemans 2017) type of knowledge. This type of knowledge doesn't directly assist governmental powers seeking to gain control of public issues. Rather, it critically reflects and appraises dominant values, interests and powers in society. Critical PA takes a 'critical distance to power' (Van Putten & Trommel 2018) and assesses how desirable existing power structures are for society with respect to the degree to which

they serve values like democracy, equality or autonomy. Van Putten and Trommel (2018) make the claim that critical PA is particularly needed now, given recently emerged forms of 'late modern pressure' (e.g. the rise of 'algorithmic behavioral control') which they view as a potential threat to liberal freedoms and democratic processes. Such forms of pressure urgently require a critical analysis, as a first step towards societal transformation.

My fieldwork ties in closely with these features of critical PA inquiry. Taking the critical camp in the behavioural state debate seriously, I have also viewed the behavioural state as a possible form of 'late modern pressure' (Van Putten & Trommel 2018). To explore that possibility, I have sought to assess the behavioural state's underlying value premises. I have empirically explored which values, interests and disciplinary paradigms are served with the rise of the emerging behavioural state. The analyses that flow from these explorations are not per se usable for behaviour experts in a direct, instrumental sense. Rather, they offer critical-reflexive insights into implicit assumptions and hidden political, discursive and power-related aspects of this community. They also reveal the backstage world of a community that has thus far been relatively hidden. Although BI has long been on the Dutch strategic policy agenda, and although some Dutch BITs have a semi-public profile (e.g. BIN NL 2017; BIG'R Website 2018), the Dutch behavioural turn has still mostly occurred out of the public eye. This stands in contrast with the Anglo-Saxon behavioural turn, which caught the attention of public media more quickly (e.g. Dunt 5 February 2014; Rutter 23 July 2015; Wintour 9 September 2010). Nonetheless, Dutch behaviour experts are placed within powerful positions and are associated with potentially impactful ideas, methods and techniques. In line with critical PA thinking, it needs to be investigated whether these powers are wielded responsibly. To do so, this study seeks to bring these experts and their practices into view more closely.

More broadly speaking, the fieldwork can also be seen as critical as it seeks to make a methodological intervention in the current behavioural state debate. Much of the arguments made herein are based on abstract ideas, ideal images, or a selection of emblematic Anglo-Saxon examples. Ethnographic research can intervene here and enrich the debate by shedding light on actual practices, also within the lesser explored 'backwater' (versus 'metropolitan') areas of the field. This type of research can help to move beyond the 'stuck' trench warfare between advocates, critics and sceptics and stimulate a more fruitful trialogue between camps instead.

Furthermore, the fieldwork is critical in its use of theory and analytical devices. I have sometimes chosen to adopt theoretical lenses even when their relevance or salience was not always immediately recognized by members of the field. For instance, I rely on the dramaturgical theory of the sociologist Erving Goffman (1959), who noted how

people behave differently in different spaces and situations. In the 'frontstage' they try to manage their public impressions. In the 'backstage', they prepare for, reflect on, and sometimes deviate from their frontstage impressions. The experts I studied would neither use the words 'backstage' and 'frontstage', nor would they acknowledge the existence thereof. I nonetheless have used this analytical device extensively. Some critical academic distance seemed needed here, knowing that the presence of a 'deviant' backstage can be denied in frontstage settings (see Brunsson (2007) on *meta-hypocrisy*). This seemed to be the case too in my field of study, where I observed differences between the 'talk' and the 'walk' that were not mentioned publicly but were acknowledged by experts in 'safe', backstage settings when prompted.

However, the fieldwork also has a more constructive side. To begin with, I have sought to familiarize myself with Dutch governmental behaviour experts, seeking to understand their world from the inside. This fieldwork, and the analysis that flows from it, is not just interesting for academic scholars. It can also increase self-understanding in the field and help behaviour experts make sense of who they are, why they do what they do, and what they are struggling with. My analysis attempts to pinpoint the behavioural state in its complexity, rather than trying to dismantle or dismiss it. I seek to debunk simplistic portrayals that overlook its diverse, hybrid and open-ended nature.

In addition to providing analyses and reflections, I have also contributed more directly and practically to the field's development. At various moments during the fieldwork, I consulted for and collaborated with behaviour experts. I was invited to several brainstorm sessions to help Dutch BITs in their search for how to further give shape to their practice. I was part of the 'Living Lab', a co-produced initiative which ran trials of public health-related behaviour change techniques in a local community building. I also helped set up the 'Urban Nudging' project, which experimented with the co-production and democratization of behavioural policies. And lastly, I was temporarily employed by a Dutch ministerial BIT, assisting the team with ongoing projects and helping to develop its core analytical tool.

***Data collection***

The empirical data for this study have been collected in four research phases over the course of four years, from September 2014 to December 2018. During these phases I used various methods to collect data. These included: holding informal conversations for orientation purposes, studying and analysing documents, doing online desk research, making informal visits to the field, conducting interviews, doing short-term participant-observations, conducting a focus group, and engaging in long-term collaboration with Dutch behaviour experts.

Phase I of the research was most exploratory in terms of the fieldwork conducted. I started by mapping and tracing activities in the Dutch BI field. I did mostly short-term fieldwork, interviewing behaviour experts and observing them for relatively short time periods. Continuously moving in and out of the field, this type of fieldwork could be described as 'yo-yo fieldwork' (Wulff 2002). To identify and select interviewees and observees, I did a preliminary mapping exercise and later on used the snowballing technique. During Phase I, my primary research locus was central government, although I also did consulting work for the local 'Urban Nudging' project. Phase II was a deliberate attempt to immerse myself more deeply into the field. From September to December 2016, I had the opportunity to temporarily work for the BIT of one of the Dutch ministries. I joined this team as an 'employee ethnographer', assisting the team in regular activities while making observations and holding conversations. In Phase III, I shifted my focus to behaviour experts at the local level, where by then I had identified increasing activity. In this phase most of the fieldwork was of a 'yo-yo' (Wulff 2002) nature. In Phase IV, I conducted a focus group with a diverse group of behaviour experts in order to validate overarching findings (see Appendix II for a list of participants and their affiliations).

Together, these four phases aggregate into an in-depth research process in which I studied the behavioural state from up close, while triangulating data sources, validating findings with both academic peers and practitioners, and offering methodological reflections (for more details see the 'Methods' sections of the individual empirical chapters). Table 1.2 provides a short summary of the research process (see also Chapter 8).

**Table 1.2: Research phases**

| *Research phase* | *Types of fieldwork* |
|---|---|
| *Phase I*<br>From 2014 to 2016 | *Mapping and exploring the field*<br>- Informal conversations with academics and practitioners for orientation purposes<br>- 24 interviews with 35 central governmental behaviour experts<br>- 55 hours of short-term participant observation at various sites<br>- Involvement as an academic adviser in the local 'Urban Nudging' project over the course of ten months<br>- Mapping and document study<br>- Informal visits to the field |
| *Phase II*<br>In 2016 | *Immersing myself more deeply*<br>- Four months of full-time participant observation as employee |

| | |
|---|---|
| | ethnographer in a ministerial BIT, plus two exit-interviews and various pre- and post-visits<br>- Document study |
| *Phase III*<br>From 2016 to 2018 | *Moving to the local*<br>- 10 interviews with 15 local governmental behaviour experts<br>- 19 hours of short-term participant observation at various sites of local governance, primarily municipalities<br>- Involvement in the local 'Living Lab' project<br>- Mapping and document study<br>- Informal visits to the field |
| *Phase IV*<br>In 2018: | *Validating overarching findings*<br>- Focus group with five behaviour experts<br>- Informal visits to the field |

***Selection strategy***

During the research process I generally used three selection criteria in defining my main study object: the 'behaviour expert'. First, behaviour experts had to be *working directly for and in Dutch government*. Second, they had to be *self-proclaimed* experts, presenting themselves as the heralds of the BI trend. Third, they had to be using behavioural insights on a *structural* and *explicit basis.* This means that I didn't study people who worked outside the public domain or used behavioural science only implicitly or incidentally. Within these criteria, I selected interviewees and observees guided by a pragmatic logic of entering the field wherever I could get access, while also seeking to get a comprehensive view on the field. At the time I started the research, most BI activities were visible in central government, which therefore provided a pragmatic starting point for research. Later on I discovered increasing BI activity in local government, which led me to shift my attention to that terrain. During the research process, I sought to capture the emerging Dutch BI landscape in its full depth and breadth as much as possible. I tried to get access to the deeper backstage areas, while also seeking to study experts in numerous sites of government – both within the innermost regions of the central state as well as within implementation bodies, regulatory agencies and municipalities.

Table 1.3 gives an overview of the institutions where ethnographic fieldwork was conducted and, when identifiable, the names of the BI configurations linked to these institutions.

**Table 1.3: Sites of fieldwork**

| | *Studied institutions and related BI configurations* |
|---|---|
| *Ministries* | - Economic Affairs and Climate Policy – *BIT EZK*<br>- Infrastructure and Water Management - *BIT IenW*<br>- General Affairs - *Leernetwerk Gedrag / Gedragslab*<br>- Ministry of Health, Welfare and Sport<br>- Ministry of the Interior and Kingdom Relations<br>- Ministry of Defence<br>- Ministry of Justice and Security<br>- Ministry of Social Affairs and Employment<br>- Interdepartmental network – *Behavioural Insights Network Netherlands (BIN NL)* |
| *Regulatory agencies* | - Dutch Healthcare Authority (NZa)<br>- De Nederlandsche Bank (DNB)<br>- Authority for Consumers & Markets (ACM) – *Kenniscentrum Gedragsverandering*<br>- Netherlands Authority for the Financial Markets (AFM) – *Werkgroep Consumentengedrag*<br>- Dutch Centre for Crime Prevention and Safety (CCV)<br>- Netherlands Food and Consumer Product Safety Authority (NVWA)<br>- Interdepartmental network - *Werkgroep Handhaving en Gedrag* |
| *Executive agencies* | - Tax and Customs Administration - *Team Gedragsverandering*<br>- Employee Insurance Agency (UWV) - *Denktank Gedragsbeinvloeding*<br>- Rijkswaterstaat<br>- Knowledge Centre for Sport Netherlands<br>- Netherlands Enterprise Agency (RVO)<br>- Netherlands Nutrition Centre<br>- National Institute for Public Health and the Environment (RIVM)<br>- The Netherlands Organization for Health Research and Development (ZonMw) |
| *Local governments* | - City of Utrecht – *Nudgenetwerk Utrecht*<br>- City of Rotterdam – *Behavioural Insights Group Rotterdam (BIG'R)*<br>- City of Amsterdam – *Gedragsdeskundige*<br>- City of The Hague<br>- City of Leeuwarden<br>- City of Hoorn – *Gedragsbeinvloeder*<br>- City of Enschede – *Adviseur Gedragsbeinvloeding*<br>- City of Deventer<br>- Divosa<br>- The Association of Netherlands Municipalities (VNG)<br>- Intermunicipal G6 network – *G6 Inzet op Gedrag*<br>- Co-production with the City of Utrecht – *Urban Nudging*<br>- Co-production with the City of Utrecht – *Living Lab* |

***Empirical context: the Dutch Behavioural Insights landscape***

This study explores the behavioural state in a particular context: Dutch government, which comes with its own unique aspects – for instance its prevailing culture of expertise or its political system. These unique aspects may well have an influence how BI is being institutionalized in the Netherlands. Therefore, this study aims to provide a contextualized analysis, exploring the 'varieties of practices of behavioural expertise' (Strassheim & Korinek 2015) that may have emerged in the 'hinterland' of the global BI landscape. Chapters 5 and 6 especially nurture this contextualizing ambition. Chapter 5 explores the developmental trajectory of Dutch BI, while linking this unique aspects of Dutch policymaking. Chapter 6 takes on a more international focus, analysing existing varieties of behavioural state practice within Dutch and Australian government.

As this study primarily concentrates on the Dutch behavioural state, let me shortly set the scene of the Dutch BI landscape. In the Netherlands, BI received attention as of the early 2010s thanks to a series of reports from several government advisory bodies, including the Council for the Environment and Infrastructure (Rli), the Netherlands Centre for Ethics and Health (CEG), the Netherlands School of Public Administration (NSOB) and the Scientific Council for Government Policy (WRR). The WRR especially played a key role in getting BI on the strategic policy agenda, having published a number of reports on the matter, starting with *De menselijke beslisser* [The human decision-maker] (Tiemeijer et al. 2009; but also see Tiemeijer 2011; Jonkers & Tiemeijer 2014; WRR 2017). In 2014, the sitting Cabinet responded to three key Dutch advisory reports, encouraging the wider use of behavioural science theory and methodology within government (Ministry of Economic Affairs 2014).

Behavioural expertise has not (yet) been deeply institutionalized within Dutch government. BI practices are developing slowly and in an exploratory manner. Since BI became part of the strategic policy agenda, various behavioural units, networks, working groups and soloists have become active. Most units possess sparse resources, and have fewer than ten members. One of the first Dutch units to emerge was the Team Gedragsverandering [Team Behaviour Change] in the Dutch Tax and Customs Administration, set up in 2009 and publicly introduced by Dutch news media (De Jong & Rusman 2 March 2015). Other pioneering units were BIT EZK of the Ministry of Economic Affairs and Climate Policy, and BIT IenW of the Ministry of Infrastructure and Water Management. The emergence of BI units has mostly occurred in a low-key, decentralized and network-like manner, with organizations experimenting with BI in their own ways, following the energy of individual enthusiasts and informal collaborations. The field's professional boundaries seem undefined and still 'in the making'.

Initially, most BI activity took place at the central state level, with expertise emerging in ministerial departments as well as in regulatory and implementation agencies. In the years to follow, BI activity also emerged at the local state level. Municipalities in larger Dutch cities (Tilburg, Eindhoven, Rotterdam, Utrecht, Den Haag, Amsterdam) established their own behavioural networks and units, and have been collaborating since in an intermunicipal G6 network ('G6 Inzet op Gedrag'). Some smaller municipalities have been developing behavioural expertise as well. Similar to central state level developments, most of the local initiatives have been organized in a low-key, ad hoc, network-like fashion – reliant on the presence of bottom-up enthusiasm. A notable outlier is Behavioural Insights Group Rotterdam (BIG'R). With over 25 members, and strong ties to the local university, BIG'R is relatively institutionalized.

As the case of BIG'R suggests, some Dutch BI practices seem to be moving from an experimenting phase to a maturing phase of development. These are mostly organizations that have already been working with BI for a number of years, have set up shop, and are now looking for next-level challenges or looking for a more structural uptake of behavioural science beyond quick wins. For instance, BIT EZK has developed its own research innovation agenda (see e.g. Ministry of Economic Affairs 2016), and BIT IenW has built its own 'DOE-MEE tool' to facilitate a broader engagement with behavioural insights. At an interdepartmental level, Behavioural Insights Network Netherlands (BIN NL) has taken on a coordinating role in facilitating and anchoring the central government's uptake of behavioural science. Since its launch in 2014, BIN NL has published an update report showcasing projects (BIN NL 2018) and practical guidelines on behavioural policymaking (BIN NL 2017b; 2018). Also, it has organized an annual event ('Dag van het Gedrag'), with keynotes and workshops dedicated to knowledge exchange in the field. The theme of the most recent event was 'Next Level', reflecting the new developmental phase that (a part of) the Dutch BI landscape seems to be heading towards.

### 1.3 Contributions to science, state and society

***Overarching contributions***

This study seeks to be relevant by studying the behavioural state from a relatively unorthodox angle. I study nudg*ers* instead of nudg*es,* and more importantly, I study them from up close. My ambition is to bring into view the behavioural state's backstage – its people and everyday practices – so as to arrive at a richer and more realistic understanding of it.

This study also starts from an atypical theoretical angle. I adopt a critical public administration perspective, introducing frameworks and ideas – for instance on professionalization or policy translation – that have thus far been unexplored in this

context. My theoretical ambition is to further expand the currently modest body of critical PA inquiry into the behavioural state.

This study contributes to the current debate by showing how and where our current understanding of the behavioural state requires a serious update. Generally, my observations reveal a need to move beyond the existent trench warfare with either utopian, dystopian or sceptic accounts which in their own ways are all simplistic. Instead, this study pushes for an understanding of the behavioural state as a complex phenomenon. Firstly, I urge scholars to understand this state from the lens of *variety and metamorphosis,* acknowledging that it is inevitably subjected to a process of policy translation as it travels across governmental institutions around the globe. The result is that varieties of behavioural state practices emerge, many of which have remained outside of the public eye. Hence, *the* behavioural state doesn't exist. Secondly, this study seeks to comprehend behavioural state practice from the lens of *friction and ambiguity,* acknowledging that behavioural units – and institutions more generally – continuously face uncertainty, constraints and conflicting demands. To cope with these, they must to some extent adapt their ideal approach. They must act hypocritically (Brunsson 2007), with a 'talk' and a 'walk' that seem to stand at odds with one another. This study tries to reveal these types of tensions and conflicts that inevitably characterize behavioural statecraft.

This study also contributes to the further development and sophistication of behavioural policy practice. While I don't push a specific agenda for BI's future development – like 'Punch!' (Jones et al. 2013), 'Steer' (Rowson 2011) or 'Nudge Plus' (John 2018) – my findings *do* help behaviour experts to see the diversity of possible practices in their own field, and their respective advantages and downsides. In pinpointing the diversity and contingency within BI practice, this study forges a space to think about the transformation thereof. In addition, this study considers what the future of the behavioural state may look like, sketching three development scenarios. These insights into how BI practice is materialized across different sites of government, and how it might evolve over time, can help behaviour experts give shape to their future course of development more consciously. It could help them adapt their practices so that they are guarded against core criticisms such as technocracy, psychocracy, and methodological narrowness.

***Contributions of the individual empirical chapters***

While all of the empirical chapters help to expose the more complex nature of the behavioural state, each individual chapter also comes with more particular results. These are summed up below.

*Chapters 3 and 4 - Practices*

Chapters 3 and 4 deal with a basic yet little explored question: is there actually a Dutch

behavioural state? These chapters touch on the deadlocked debate about the behavioural state's actual prevalence. While sceptics tend to frame the behavioural turn as a mere rhetorical fad, advocates continue to promise that BI is 'here to stay'. Chapters 3 and 4, however, present a more nuanced picture of a modestly present, but still institutionalizing landscape. Chapter 3 explores actual practices of behaviour experts at the central state level. It uncovers how these experts are best understood not as intervening choice architects but as 'knowledge brokers', ensuring that relevant knowledge flows smoothly between and within organizations.

Chapter 4 explores actual practices at the local level – a possibly particularly fruitful BI landscape given that local officials tend to deal with many concrete and localized behavioural challenges. The chapter shows a moderate local behavioural turn, with behaviour experts beginning to explore the opportunities of BI. Their further institutionalization depends on their ability to adapt to their local context, and deal with its constraints and competing demands.

*Chapters 5 and 6 - Professionalization*
Chapters 5 and 6 concentrate on the overarching community of BI. How can we depict its developmental trajectory thus far? Drawing upon the professionalism literature, Chapter 5 traces how BI has been institutionalized in Dutch central government, particularly paying attention to what unifies or divides the field. It shows how behavioural expertise has thus far developed in a loosely fragmented fashion. Behaviour experts possess a basic sense of group identity, but beyond they draw from diverging methods, ideas and techniques.

Chapter 6 also explores how the BI field has been developing, although here the focus lies on international developments. Ethnographic fieldwork on both Dutch and Australian practices are combined order to distil key areas of consistency and contingency in the international BI landscape. The chapter shows how the field is subject to a dynamic of policy translation, in which individual behaviour experts are adapting BI's espoused agenda to fit their own local context. Particularly deep (and puzzling) adaptation exists at the methodological level.

*Chapters 7, 8 and 9 - Expertise*
Chapters 7, 8 and 9 look at the role of expertise within the behavioural state. They try to uncover how this state 'sees'. These chapters draw on the PA literature on the multifaceted role of knowledge in the policy process. This literature helps to see why behavioural turn within policy is far from guaranteed and why the professed epistemic gaze of behaviour experts can be difficult to apply in practice. More specifically, Chapter 7 explores the institutional logics that inform behavioural state practice. BI's role models seem to promote a neorationalist logic, treating policy design as a purely managerial and scientific affair. This portrayal clashes with the

hitherto more incrementalist strategy of policymaking, in which policy decisions flow from a chaotic coming-together of all sorts of more-than-rational factors (including inertia, ideological conflict and bureaucratic turf wars). Chapter 7 analyses how behaviour experts negotiate these conflicting institutional logics and come to follow a more hybrid, 'rationalized incrementalist' kind of strategy in the backstage.

Chapter 8 looks at the role of expertise in the behavioural state in relation to two major criticisms in the scholarly debate. Critics have associated the behavioural state with elements of technocracy, and also 'psychocracy' (Jones et al. 2013), a policy style that relies exclusively on psychological theories and methods. Chapter 8 scrutinizes whether these criticisms hold up in practice. It argues that these critiques are sometimes but not always valid, as they fail to reflect some alternative developments in the field that do respond to these critiques.

Chapter 9 explores the role of interdisciplinarity in the behavioural state. Drawing from insights both from the field of PA and human geography, it analyses the thus far bounded forms of interdisciplinarity within the BI landscape. The kind of interdisciplinarity that exists in the field is mostly confined to the fusion between psychology and economics. At the same time, the chapter also points out early signs of more expansive forms of interdisciplinarity within the field.

Before we move to the empirical chapters, the chapter that follows will outline a brief history of the contemporary behavioural state. When and why did the BI movement arrive on the global policy scene and how has it spread since then? Subsequently, the chapter reviews the current behavioural state debate. Who are the leading voices and what are their main claims? Which concepts, ideas and values are central in the debate? As these questions are further explored, the behavioural state is situated in its wider intellectual, political and institutional context.

*Chapter 2*

# Debating the Behavioural State

## 2.1 The rise of the Behavioural Insights movement

The starting point for this study is the recent emergence of the Behavioural Insights movement onto the global policy scene. This movement has as its main aim the uptake of a body of so-called 'behavioural insights' in policy. Its most direct origins lie in the relatively recent birth of the new behavioural economics school (hereafter: NBE), and the popularization thereof in the public realm. This NBE school tries to understand human behaviour by unravelling the individual cognitive decision-making processes that drive them. It relies on a fusion of cognitive psychology, cognitive design, some strands of neuroscience, and economic thinking (Whitehead et al. 2017; Sent 2004). Although NBE emerged in the early 1970s, it is only within the last two decades that it seems to have caught the popular imagination. By now, as any international airport bookshop will tell, the list of NBE bestsellers is vast and includes books like *Thinking, Fast and Slow* (Kahneman 2011), *Predictably Irrational* (Ariely 2008), *Blink* (Gladwell 2005), *Animal Spirits* (Akerlof and Shiller 2009), and *The Small Big* (Martin et al. 2014). Thaler and Sunstein's *Nudge* (2008) has perhaps been the most influential book in the genre. The Nobel Memorial Prizes awarded to two of NBE's thought leaders – to Daniel Kahneman in 2002 and to Richard Thaler in 2017 – have further amplified the discipline's public profile and prestige.

NBE has also worked its way into the world of public policy. NBE-minded 'Behaviour Change' agendas first emerged in the United Kingdom and the United States. In 2004, the UK Cabinet Office published a paper on 'Personal Responsibility and Changing Behaviour' (Halpern et al. 2004), noting how NBE insights might be translated into innovative policy strategies. During the decade thereafter, NBE-style policy thought gained significant traction within Anglo-Saxon governments and increasingly also beyond the Anglosphere. One sign of NBE's growing popularity was the increasing acknowledgement from think tanks and advisory bodies regarding the potential of behavioural insights to produce more realistic and effective public policy (e.g. Dolan et al. 2010; Haynes et al. 2012; Service et al. 2014; Tiemeijer et al. 2009; Jonkers & Tiemeijer 2014). Additionally, leading NBE scholars were increasingly consulted by governments for their valued expertise. For instance, in the UK, *Nudge* author Richard Thaler would regularly act as adviser to Premier David Cameron, while in the USA his co-author Cass Sunstein was installed by President Barack Obama as 'regulatory czar' at the White House's Office of Information and Regulatory Affairs (Halpern 2015).

A key event in the history of BI was the founding of the first 'Behavioural Insights Team' ('BIT' in common parlance). The trendsetting BIT UK was launched in 2010 by Prime Minister David Cameron as a central strategy unit within the Cabinet Office. Since then, it ran projects in a wide variety of policy domains (e.g. relating to debt,

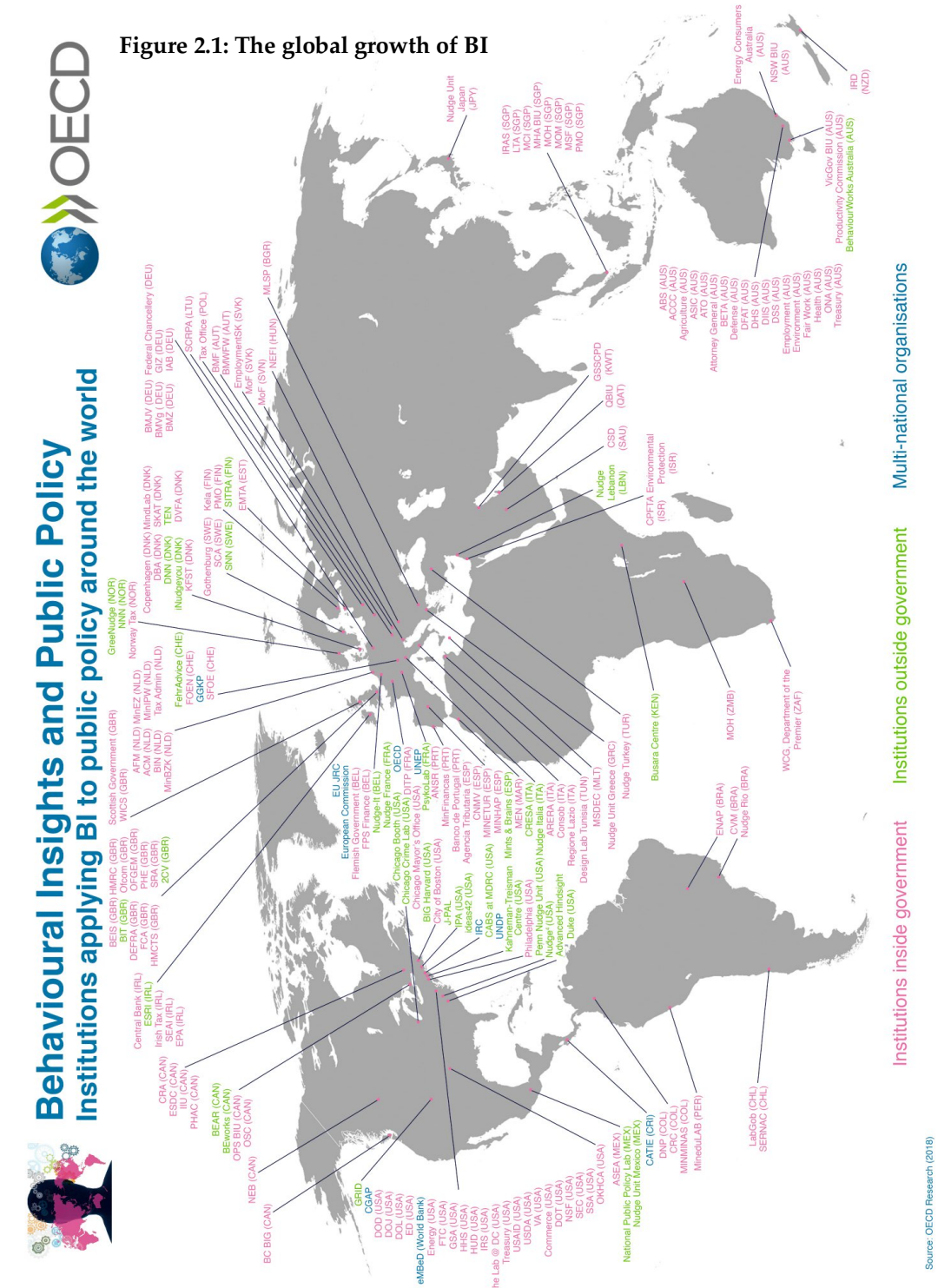
health, tax, charity, and sustainability), typically designing experiments to test the effects of NBE-based policy solutions, mostly nudges. One of BIT UK's best-known success stories is that it ran a large-scale trial involving 100.000 tax payers, telling them that most of their neighbour's *had* already paid their taxes. Effectively exploiting people's tendency to 'follow the herd', this social feedback nudge led to up to a 5% increase in payments and allegedly helped the Cabinet Office bring in 30 million pounds a year in income tax (Service et al. 2014; Rutter 23 July 2015).

Following in BIT UK's footsteps, many international policy actors launched similar initiatives of their own, including the European Commission, the OECD, the World Bank, and governments in France, Germany, Belgium, Canada, Qatar, Ireland, Lebanon, Denmark, Australia, Japan, the Netherlands, the United States, and Singapore (e.g. Lunn 2012; Van Bavel et al. 2013; Lourenço et al. 2016; Jones et al. 2013; Whitehead et al. 2014; 2017). A worldwide BI community was in the making. Numerous BITs were launched, including BIN NL, Behavioural Economics Team of the Australian Government (BETA), Nudge Lebanon, Qatar Behavioural Insights Unit and Nudge Lab Pakistan. This behavioural turn took place both in and outside of the public sector (for more detail on the corporate BI landscape, see Whitehead et al. 2017, Chapter 7 on 'The Neuroliberal corporation'). Beyond the work of these specialized units embedded in or near the core executive, BI initiatives were also undertaken by governments in a more temporal, ad hoc, project-based fashion. The OECD (2017, see @faisal_naru 2018) has recently mapped the global spread of institutions making use of behavioural insights. This map is presented in Figure 2.1.

As often noted, the rise of BI is only the latest manifestation of the behavioural state. Governments have form when it comes to embracing the modernization, scientization, and more particularly, 'behaviouralization' or 'psychologization' of policymaking. Undeniably however, BI's ascendancy has given a new impulse to behavioural expertise in policymaking – as clearly illustrated by Figure 2.1 (OECD 2017; see @faisal_naru 2018). Distinctive about this latest impulse is that it carries a specific strand of behavioural science: NBE, earmarked by its rapidly developing catalogue of biases in human decision-making (for an overview see Benson 2016).

In this chapter I situate the BI movement in its broader historical, intellectual and political perspective. I do so by examining the considerable debate that BI's emergence has sparked, and shedding light on the field's academic and institutional legacies.

**Figure 2.1: The global growth of BI**



## 2.2 Trench warfare about the behavioural state

The rise of BI has set off a compounded debate. While primarily taking place within academic and political-administrative circles, this debate has also reached the level of broader public discourse (*inter alia,* Rutter 23 July 2015; Wallace-Wells 13 May 2010; De Jong & Rusman 2 March 2015; Easton 11 October 2016). As already pointed out in Chapter 1, the emerging behavioural state has been appraised in widely diverging ways. Following Towfigh and Traxler (2016) in their claim that 'Nudges Polarize!', one could even argue that the debate has become a 'trench warfare' between radically opposed camps (Towfigh & Traxler 2016). We can broadly distinguish between three 'trenches': advocates, critics and sceptics.[1] This chapter continues by throwing light on these three particular trenches.

***Advocates***

The BI movement has the backing of a growing camp of agents who stress the great promise that Nudge and BI holds for public policy (*inter alia,* Dolan et al. 2010; Thaler and Sunstein 2008; World Bank 2015; OECD 2017; Lourenço et al. 2016; Shankar & Foster 2017). This camp is championed by renowned academics (e.g. Daniel Kahneman, Richard Thaler and Cass Sunstein), but also influential think tanks and consultancies (e.g. ideas42, Dijksterhuis & Van Baaren and Nudge Unit Greece), as well as transnational institutions (e.g. The World Bank, the United Nations, the European Commission and the OECD) and specialized BITs (e.g. BIT UK, Behavioural Insights Unit New South Wales, BIN NL). Governments' recent behavioural turn is also supported by the launch of new international academic journals, including the Journal of Behavioural Public Administration and the Journal of Behavioural Public Policy. According to all of these advocates, BI represents an innovative and much needed update to what in an encompassing sense could be taken as the hegemonic policy agenda of New Public Management (Esmark 2019).

New Public Management (hereafter NPM) emerged in western liberal democracies since the early 1970s and became dominant from the 1980s onwards. Firmly rooted in neoliberal and managerialist schools of thought (Hood 1991; Dunleavy & Hood 1994),

[1] It should be recognized however that some overlap exists between these three 'trenches'. Both critics and sceptics for instance share that they dismiss the behavioural state, although they do so for wholly different reasons. Further, some scholars could easily be divided in multiple camps. For instance, Whitehead et al. (2017) are both broadly sympathetic of the emerging BI movement and yet critical about some aspects of it. Oliver (2015), while a staunch advocate of the use of NBE in policymaking, is sceptic about the value of BI associated terms such as Nudge and Libertarian Paternalism. Despite these partial forms of overlap, the distinction between three trenches nonetheless still provides a basic framework that bring into view the diversity of viewpoints in the debate.

its primary ambition was to achieve greater levels of policy effectiveness and efficiency. While previous policymaking styles relied strongly on the classical judicial perspective, NPM was heavily influenced by a neoclassical and new institutionalist economic perspective. At the heart of this perspective lay the idea of the self-interested consumer with unbounded rationality and unreserved self-discipline – the *homo economicus.* This 'ideological dominance of economic norms' (Christensen and Lagreid 2002: 268) within NPM was concretized by means of a series of iconic 'NPM doctrines' (see Hood 1991). These doctrines favoured the neoliberal design logics of incentivization, disaggregation, privatization, competition and marketization as the preferred routes towards a more effective and efficient public service delivery (Hood 1991; Christensen and Lagreid 2002; Dunleavy et al. 2006).

However, in the decades to come after NPM had emerged in the 1970s, the limits, complexities and tensions associated with NPM would become increasingly clear (e.g. Davies 2014; Dunleavy et al. 2006; Hood & Peters 2004; Trommel 2018). Its grand initial promises about increased performance and efficiency were not, or only partially fulfilled, or gave rise to new types of market failure. Various iconic NPM doctrines were called into question and downplayed, for instance the 'great expectations' about the privatization of public services as a means towards improved government performance (Overman 2016). The diminishing hope in these doctrines was essentially the result of an overarching intellectual flaw within the NPM project: its weddedness to the *homo economicus,* i.e. the assumed predominance of perfectly rational decision-making in the social sphere. This assumption of humans as perfectly rational utility maximizers turned out to be at odds with the emerging picture from the behavioural sciences showing the paramountcy of extrarational decision-making instead (e.g. Tversky & Kahneman 1974). Against the background of this elemental flaw within NPM, BI advocates have legitimized their approach and forged their narrative as a 'post-NPM' megatrend (Esmark 2019). As the hope in NPM paled, they were able to arouse a new hope in improved government performance, based on a distinct – both epistemologically and morally motivated – advocacy. What that advocacy entails, is discussed below.

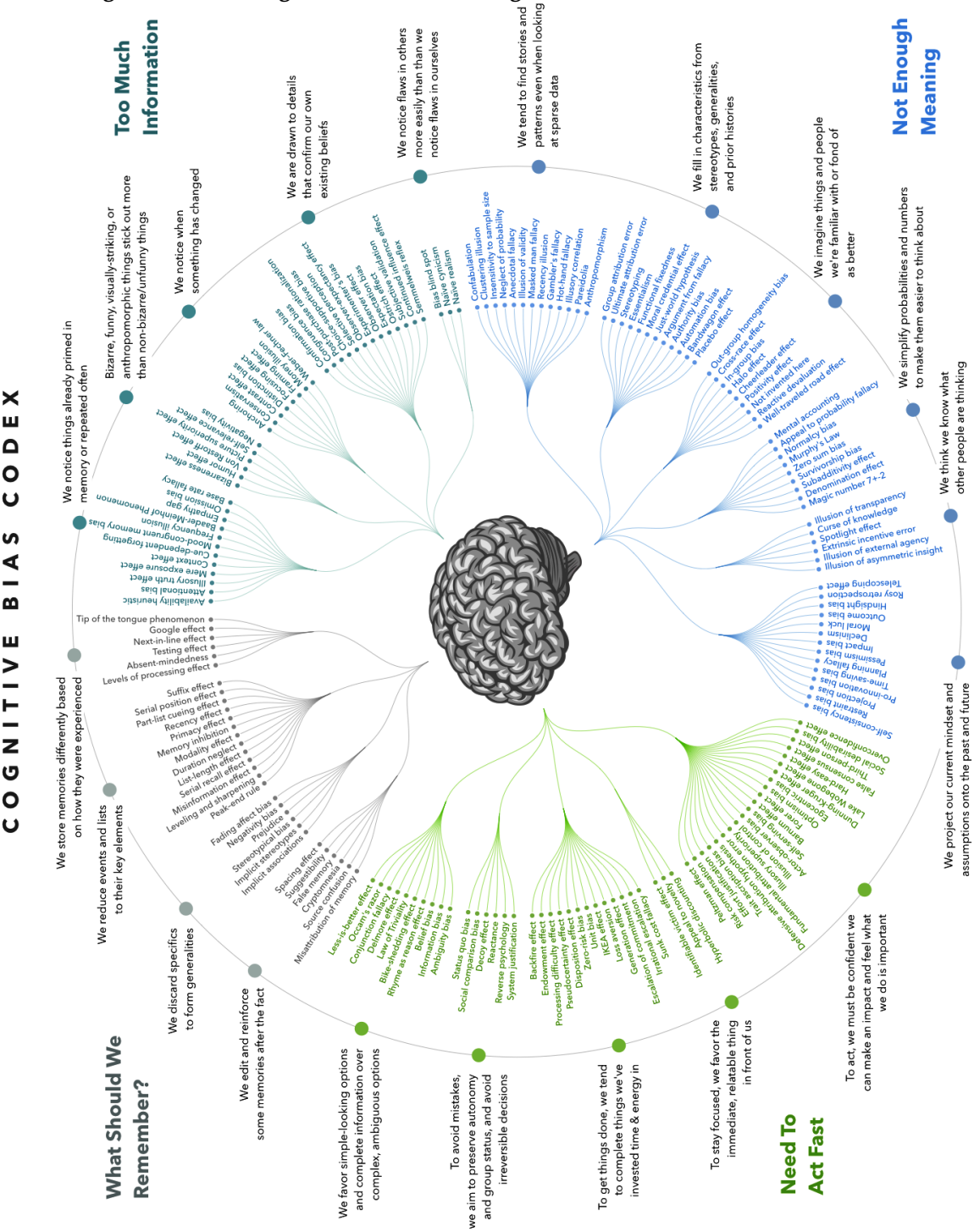
*Epistemological advocacy*

In the accounts of BI advocates, the history of their field usually starts with the discovery of in their eyes paradigm-changing insights coming from the NBE school. This is not to say that this is the objective starting point of the field but rather a perceived/narrated starting point. Alternative historical narratives exist too, offered particularly by the sceptical 'trench', pointing for instance at the preceding work of Herbert Simon (1947; 1985) and his *old school of behavioural economics*. But for the advocates however, the field's intellectual story tends to begin with NBE. This NBE school emerged since the 1970s, with a leading role for the economist Richard Thaler,

based at the University of Chicago. Thaler, as he kept a list of 'odd' non-rational economic behaviours, developed an interest of how seemingly irrelevant factors could influence economic decision making. This interest led him to the field of cognitive psychology. One of NBE's founding moments was Thaler's discovery of the work of cognitive psychologists Daniel Kahneman and Amos Tversky, who had been cataloguing the systematic errors in human decision-making since the late sixties (e.g. Tversky & Kahneman 1974). Thaler found that cognitive psychology offered innovative explanations to economists explaining why these errors occurred, linking these errors to the fact that humans make decisions in the context of uncertainty, limited time and limited cognitive capacities. He realized that if the discipline of economics wanted to be more sensitive to real-world human behaviour, it would need to engage in an interdisciplinary dialogue with the cognitive psychological sciences. It is this attempted interdisciplinary dialogue – initiated by Thaler and later continued through fellow behavioural economists like Colin Camerer, George Akerlof, Iris Bohnet, Dan Ariely, George Loewenstein and Eldar Shafir – which defines NBE.

With as their general mission to incorporate cognitive psychology into economics, the more detailed work of NBE scholars consisted in cataloguing the diverse patterns in which humans deviate from rational decision-making. They found out that humans are 'heuristic' reasoners: they follow all sorts of mental shortcuts ('heuristics') that help them choose efficiently in a context of limited time, knowledge and cognitive bandwidth. These heuristics, although crucial in using one's mental processing capacities parsimoniously and efficiently, can however also generate systematic deviations from rationality, also annotated as 'flaws', 'errors' or, in more formal NBE terminology, 'biases' in decision-making. Starting with Thaler's provisional list of 'odd' economic behaviours, NBE scholars have since then identified a massive body of biases, including *status quo bias, authority bias, stereotypical bias, hyperbolic discounting, loss aversion, bias blind spot, confirmation bias, peak-end rule attentional bias, optimism bias,* and also some with more intriguing names such as *the cheerleader effect, Maslow's hammer,* and *the tip of the tongue phenomenon* (Benson 2016). Benson (2016) has recently categorized and mapped out the many different biases (175 in total according to his catalogue), presented in Figure 2.2.

**Figure 2.2: A catalogue of decision-making biases (Benson 2016)**



As these various biases touch on numerous, widely diverging aspects of human decision-making, going into detail about each of them would go beyond the scope of this study. Nonetheless, it could be said that NBE studies into human decision-making biases reflect three basic insights. First, humans are rationally bounded. They heavily rely on mental shortcuts to make daily decisions in a context of limited time, information and cognitive bandwidth. Second, humans have limited self-control. Third, they are easily influenced by cues within their surrounding social and material environments (BIN NL 2017a; Jonkers & Tiemeijer 2014). A most efficient and influential attempt to summarize the discoveries of NBE scholars has been made by Daniel Kahneman. In 2011 he published a synthesis – titled *Thinking, Fast and Slow* – of the work that he and colleagues had done over the last four decades on the 'quirks' of human decision-making. This book quickly became an international bestseller and a foundational work for experts in the BI field. The simple yet strong contribution of this book was its pushing of a parsimonious theoretical framework, basically consisting of two, interacting, cognitive systems used in human decision-making – *System I* and *System II*. As Kahneman explains, *System II* refers to the slow, calculative and conscious thinking processes associated with the brain's prefrontal cortex, while *System I* concerns the fast, intuitive and automatic thinking processes associated with parts of the brain that have older evolutionary origins. Table 2.1 (based on Dolan et al. 2010; Kahneman 2011) further clarifies how these systems differ from one another. This 'simple' juxtaposition between *System I* and *II* has further helped NBE scholars challenge mainstream economists' holy cow – the assumption of the *homo economicus*. It helped making the idea of bounded rationality in human decision-making more intuitively intelligible.

**Table 2.1: System I and II thinking**

| *System I: Reflective* | *System II: Automatic* |
|---|---|
| Controlled | Uncontrolled |
| Effortful | Effortless |
| Deductive | Emotional |
| Slow | Fast |
| Self-aware | Unconscious |

The new insights from NBE, parsimoniously captured by the model of *System I* and *II,* had important implications for policy development. They provided a deep explanation for why NPM's promise of improved government performance had not been fully met. As long as policies would continue to adopt NPM's underlying assumption of the citizen as *homo economicus,* they would struggle to solve policy issues – whether these related to personal finance, mobility, sustainability, safety, consumer protection, public health, or other areas. Consider obesity policy, as an

example. Public domains have increasingly become 'obesogenic environments' with damaging impacts on the well-being of individuals (Lake & Townshend 2006). The behavioural impacts of these environments however have not been thoroughly addressed within the NPM paradigm, assuming that unhealthy choices are made freely and rationally.

Against the background of the 'behavioural market failures' (Sunstein 2014) perceived within NPM, BI advocates have been calling for the development of a new mode of policymaking. A mode that, instead of assuming citizens as predominantly *System II* driven reasoners, is more sensitive to their *System I* decision-making. Particularly, these advocates have been promoting Nudge theory. Nudge theory, introduced by University of Chicago economist Richard Thaler and Harvard law scholar Cass Sunstein, refers to the redesigning of choice environments in subtle ways that go with the grain of *System I* behaviour. Nudge theory is firmly rooted in NBE insights into the *systematic* deviations from rational decision-making. This systematic nature is important for Nudge, because in revealing humans as '*predictably* irrational' (Ariely 2008; italics JF), their irrationality also became measurable and manageable. Nudge theory capitalizes on this – in the context of policymaking very promising – line of thought. It translates NBE insights into human irrationality into individual and social welfare promoting behaviour change interventions.

Nudge theory doesn't just start from, and justify itself on the basis of, the observation of a dominant *System I*. It also aims to *tap into* aspects of this *System I*, as a route towards steering citizens towards desired behaviours. It does so by subtly redesigning the 'choice architectures' (Thaler & Sunstein 2008) that surround people. These architectures contain all sorts of informational or physical elements that while initially seeming irrelevant have since the birth of NBE actually been (re)discovered as having substantial impacts on how citizens behave. The choices we make for instance depend on what is proximate, visible, available, attractive, socially valued or defaulted (Thaler & Sunstein 2008; Service et al. 2014). Importantly, Nudge proponents recurrently note, such contextual elements influencing behaviour cannot be avoided, as people are always surrounded by some kind of context that must favour certain options over others (for instance in terms of their proximity). Nudge theory embraces this inevitably contextually predisposed nature of public spaces, and from there seeks to design these spaces more consciously with individual or social welfare in mind (Thaler and Sunstein 2008).

Nudge theory comes with its own special toolkit, featuring various types of 'nudges'. These nudges are typically juxtaposed with the classic triad of policy instruments consisting of 'sticks' (law and regulation), 'carrots' (financial incentives), and 'sermons' (education) (Bemelmans-Videc et al. 2003). Nudges, unlike sticks, don't

restrict or enforce choices. They merely restructure and reframe choice options, while always leaving room for recipients to 'opt out' with minimal costs. Nudges, unlike carrots, don't employ financial incentives. Rather, they use the type of contextual, non-financial incentives identified by NBE scholars. Nudges, unlike sermons, don't try to rationally persuade people. Instead, they seek to evoke behaviour change trough addressing the 'more-than-rational' nature of human decision-making.

The intervention techniques that sit within Nudge's toolkit are fairly diverse. They include: default changes; feedback; warnings; action plans; simplifications; and changes in the physical environment (for other typologies see, *inter alia,* Thaler & Sunstein 2008; Van Oorschot 2013; Sunstein 2014; Hansen & Jespersen 2013). Specific examples of nudges are even more diverse and come in the dozens. A classic nudge is the 'Piano Stairs', visualizing the treads of a staircase as the keys of a piano. This nudge draws on the psychologic mechanism of attractiveness to stimulate the healthy use of stairs (versus lifts or escalators). Another classic nudge is 'Save More Tomorrow': a future-minded pension plan that asks employees to commit to gradually increasing their pension saving at a later time in the future, helping them to make financially responsible decisions while avoiding the experience of a sense of loss in the here-and-now (Thaler & Sunstein 2008). A third classic nudge is the default change – from 'opt-in' to 'opt-out' – in organ donor registration policy. This simple switch, tapping into the psychology of human inertia and the lure of the status quo, can lead to up to 99% donor registration rates in some countries (Johnson & Goldstein 2003). This nudge demonstrates the emblematic promise of Nudge: small changes, big effects (although see Chapter 8 for a more critical reflection on the effectiveness of this nudge policy).

Besides the promise of innovative and more realistically underpinned policy, BI advocates also note that their approach increases the methodological rigour of governments. Acknowledging that human behaviour is complex and context-dependent, they push for the case-by-case empirical evaluation of policies in order to find out 'what works'. More specifically, they push for the use of Randomized Controlled Trial (hereafter RCT) methodology. An RCT is a field experimental design that tests the effects of an intervention on randomly selected control and treatment groups. Through these design logics of establishing control groups and randomization, an RCT can demonstrate a causal link between interventions and outcomes, while minimizing biased estimates of these outcomes for instance because of selection bias (Haynes et al. 2012).

BIT UK particularly has been a forerunner when it comes to the promotion of RCTs as an essential part of BI. This unit wrote a special report – *Test, Learn, and Adapt* – to specify and promote its own methodological approach (Haynes et al. 2012). The

report links the complexity of how humans really respond to policies to the necessity of RCTs as a most powerful way to find out 'what works'. These RCTs tend to generate novel and unexpected insights as

> confident predictions about policy made by experts often turn out to be incorrect. RCTs have demonstrated that interventions which were designed to be effective were in fact not (...) They have also shown that interventions about which there was initial scepticism were ultimately worthwhile (Haynes et al. 2012: 15)

The BI advocates' push for RCTs ties in closely with a wider trend of *evidence-based policymaking* (hereafter: EBP). EBP emerged since the 1990s in the Anglo-Saxon policy domain, where it was inextricably linked with the idea of a professionalizing and modernizing government which organized itself along the principles of being accountable and showing 'what works' (Parsons 2002; Clarence 2002). EBP strives after a strongly science-driven type of policy analysis, based on thorough fact-finding and rigorous methods (Sanderson 2002; Davies & Nutley 2000; Head 2008; Cairney 2016). As RCTs, given their rigour and explanatory power, fit perfectly within such an ambition, the knowledge derived from such trials stands highest in EBP's hierarchy of evidence.

*Moral advocacy*

Moving away from epistemological arguments, BI advocates also defend their approach from a moral point of view. They claim that Nudge is less obtrusive than traditional variants, as it doesn't restrict or enforce choices. In addition, it doesn't alter individuals themselves but merely their surrounding 'choice architectures' (Thaler & Sunstein 2008; Whitehead et al. 2017). In the advocates' eyes, their contextual approach overcomes a quintessential moral conundrum of modern states. On the one hand, states uphold utilitarian, behaviour change oriented values. They formulate policy ambitions for a 'better' society and develop instrumental strategies to meet those ambitions. On the other hand, states possess liberal, 'enlightened' values. They believe in the humanist ideal of 'the free citizen' that has dominated the western political climate of the last two centuries (Davies 1997). These utilitarian and liberal values seem to be at odds with one another. While utilitarian values ask states to tighten their control over the citizen, liberal values would ask them to ease their control (although for a more sophisticated view that takes into account different understandings of freedom, see Vugts et al. 2018). As BI advocates see it, NPM has been unable to negotiate these fundamental value tensions. It overly prioritized the value of individual choice, overlooking the rationally bounded nature thereof, and as a result thereof underperformed on its utilitarian ambitions.

Fortunately, BI advocates argue, Nudge provides a new way to renegotiate governments' behavioural and liberal values (Whitehead et al. 2017). With Nudge, governments can actually have their cake and eat it: they can shape public behaviours without substantially infringing on people's liberal freedoms. Nudges merely restructure or reframe existing choice environments, and they always ensure the possibility to 'opt out' easily (Thaler & Sunstein 2008). Moreover, nudges ideally only steer people towards goals that people themselves subscribe to (Sunstein 2017b). In this sense, Nudge forms a politically attractive middle ground between laissez-faire politics and active state paternalism. It satisfies both governments' paternalistic instinct to protect citizens from harming themselves, as well as their liberal instinct to preserve individual choice and not seek to interfere with the private life and personal values of citizens. This is why Nudge was originally introduced as 'Libertarian Paternalism' (Sunstein & Thaler 2003), fusing together two seemingly antithetical political visions on state intervention.

At a deeper level, BI advocates' advocacy of Nudge reflects their attempt to renegotiate the social state-citizen contract. They seek to re-evaluate the value of individual choice in late modern societies. From their perspective, safeguarding or further expanding individual choice is not necessarily beneficial for citizens (Jonkers & Tiemeijer 2014). Modern societies have become increasingly complex; citizens must choose more often and between more options than past generations. Also, major life choices have become more uncertain as the risks behind them (e.g. concerning mortgages, insurances, and pensions) are less clear in today's globalizing society (Van Staveren et al. 2014). In this light, terms as 'choice pressure', 'choice stress', 'choice overload' and 'paradox of choice' (Schwartz 2004) have been coined to refer to the burden that citizens can experience when forced to choose. BI advocates observe many situations where citizens are confronted with too many choices, where choices are too complex or where there are too many putatively 'bad' options. In these situations, BI advocates claim, giving people more freedom from state influence can paradoxically curtail their freedom to live their life in accordance with their own values (see Berlin 1958 on the tensions between 'positive' and 'negative' liberties). Hence, more fundamentally, BI advocates want us to rethink and (slightly) temper governments' deep liberal ambition to safeguard individual choice.

BI advocates have not just vouched for their approach based on their own moral compass. They have found extra support for their approach in an emerging body of empirical research (partly conducted by themselves) into the public acceptability of Nudge and Libertarian Paternalist approaches. Several studies have thus far showed a public majority approval for nudge type of policies, especially in the case of 'educative' nudges (e.g. Sunstein et al. 2018; Reisch 2017).

*The way forward*

To summarize, BI advocates have championed their approach against the background of a diminishing hope in New Public Management. Unlike NPM, BI policy would be more realistically underpinned with the rationally bounded citizen in mind, more rigorously evaluated, and thus more likely to be effective. Furthermore, BI policy tends to be relatively cheap and unobtrusive, making its promise complete: regained state control over even the most 'wicked' social issues without hurting the government budget nor the liberal ideal of 'the free citizen'.

For advocates, the way forward is to keep expanding the BI field along the same lines. Some small methodological tweaks or ethical preconditions may be needed, but in principle the idea is to further roll out the behavioural state. Advocates stress that the BI movement has already come a long way, and has already achieved substantive results in promoting individual and societal welfare. At the same time, there are still plenty of new topics to study and issues to tackle. In the recently launched Journal of Behavioural Public Policy, Sanders et al. (2018) reflect for instance on where opportunities for the field lie, mentioning among other things 'nudging organizations' and 'scaling interventions'. And Sunstein (2017b: 65) states that 'Money has been saved; so have lives. There is much more to do, and much more to learn.' Last, in his *Inside The Nudge Unit*, David Halpern (2015: 350-351), the director of BIT UK, notes that his unit's

> influence has already been remarkable. Policy changes driven by BIT and its sister units have led, and are leading, to millions of healthy life years saved, hundreds of thousands getting into work faster, and millions in revenue being brought forward ... one thing is sure: nudging – the use of behavioural insights and the experimental methods it has brought in its wake – are here to stay.

***Critics***

Despite the widespread popularity of the BI movement, especially within policy circles, its rise has not been appraised merely positively in the broader debate. A second, more critical trench has emerged as well. Analogue to BI's advocacy, BI's criticism can be broken down into epistemological and moral claims. These claims will be spelled out below.

*Epistemological criticism*

Part of the criticism directed at the behavioural state relates to epistemology. Critics challenge whether BI's methodological approach is comprehensive and rich enough, and whether it produces (sufficiently) effective outcomes that stretch beyond the short term. To start with concerns about effectiveness, Marteau et al. (2011: 265) note that the current evidence base for nudging reflects both an 'absence of evidence as

well as evidence of little or no effect.' Very little systematic and meta-analytical research exists into how effective nudges really are. And insofar effectiveness is proven, the effect size of nudges may still be disappointing. A repeated claim is that mere nudge approaches will not be enough to solve thorny, enduring policy issues. More heavy-handed, regulatory approaches are often needed (e.g. Marteau et al. 2011; Goodwin 2012). For instance, the House of Lords Science and Technology Committee (2011) in the United Kingdom concludes that the use of nudges in isolation is less effective. Despite their popularity in British government, the committee encourages the development of broader policy mixes. In *Against Autonomy,* Conly (2012) comes to a similar conclusion, based on the same NBE insights that BI advocates rely on. Conly argues that insights into irrational behaviour don't call for at most mildly effective nudge approaches but rather for more coercive, and thus plausibly more effective, approaches. De Jonge et al. (2018) make a somewhat different argument, claiming that nudge policies may actually prove counter-effective. This is because they can trigger 'reactance' behaviour in nudged policy recipients, particularly when they feel misled or manipulated by an untrusted source.

Besides concerns about BI policy's effects (or rather, the lack or unexpected sides thereof), critics also voice concern about BI's epistemological underpinnings. One worry is that this approach is too narrow. A number of critics assert that BI is 'undersocialized'. It focuses solely on the individual, cognitive psychological determinants of policy-related behaviours, while overlooking the wider socio-economic, institutional and historical determinants (Jones et al. 2013; Bonell et al. 2011a; Shove et al. 2012; Shove 2010; Whitehead et al. 2017). This neglect of the wider structuralist policy perspective turns the state into a 'psychocracy' (Jones et al. 2013) that can only analyse and intervene from an individualist psychological perspective. Adding to the argument, Mols et al. (2015) in their paper 'Nudge Is Not Enough' note how Nudge theory is narrowly grounded in the social cognition strands within cognitive psychology. As these strands are strongly fixated on individual cognitive processes, they tend to disregard the behavioural impacts of non-individual aspects such as social identity and group culture. Fitzpatrick (2011) questions BI's sole emphasis on individual behaviour change in the first place. He claims that the role of the behavioural component to policy issues is largely overstated.

In addition, Rowson (2011) argues that BI's approach is too shallow. It remains too much at the surface by merely seeking to change behaviours, and not underlying attitudes and values. He states that BI offers mere 'technical solutions' to what in fact often are 'adaptive challenges', i.e. complex issues that require value transformation and bottom-up participation. Mols et al. (2015) connect this shallowness with further concerns about the effectiveness of BI policies. They argue that Nudge works through superficial and passive forms of 'norm adaptation', not addressing people's values

but merely their behaviours in particular contexts. However, a more active form of 'norm internalization' – one that appeals to people's beliefs and identity – would be needed to achieve lasting behaviour change (Mols et al. 2015).

In a broader sense, critics have challenged BI on its alleged naïve rationalism. They object to its rationalist portrayal of policymaking, assuming a linear and instrumental relationship between science and policy (Shove 2010). Such a portrayal overlooks that policy decisions are generally made in a context of inherent uncertainty and a plurality of viewpoints. The policy world is marked by 'VUCA' (i.e. *v*olatility, *u*ncertainty, *c*omplexity and *a*mbiguity; Van der Wal 2017). Knowledge is often absent, inconclusive, and/or contested. Also, it can be used for not instrumental but symbolical, political or irrational reasons (Ingold & Monaghan 2016; Cairney 2016; Weiss 1979). Hence, BI's emphasis on the instrumental usage of knowledge in policy, captured by its adage of 'applying behavioural insights', is unrealistic. Lodge and Wegrich (2016) aptly observe a 'rationality paradox' in BI. While it is acutely aware of the bounded rationality of citizens, it tends to overlook the many forms of bounded rationality of governmental processes and producers. It ignores that policy decisions are, among other factors, shaped by happenstance, satisficing, inertia, political struggle, bureaucratic turf wars and the appeal of the path of least resistance (e.g. Lindblom 1959). Such factors are likely to thwart BI's attempts at (fully) rationalizing and instrumentalizing the policy process. Lastly, coming more broadly from the critical literature on EBP, concern is also voiced regarding BI's particular hierarchy of evidence, in which RCTs are taken as the 'gold standard' of evidence. The criticism is that BI wrongfully assumes that RCTs can provide fixed and universal knowledge about 'what works'. Rather, RCTs provide isolated knowledge about 'what worked' (Biesta 2007) in a particular context. Moreover, the extreme focus on RCTs in BI may lead to the exclusion of a range of alternative, softer methods which also produce uniquely valuable types of knowledge. For instance, interviews can provide qualitative understandings into the beliefs and behavioural drivers of target groups. And building professional work experience can provide essential contextual knowledge about how best to implement interventions locally (Parsons 2002; also see Chapters 3 and 6).

*Moral criticism*

Critics have also voiced moral concerns with BI. An initial core concern holds that the rise of the behavioural state means a stark increase in covert and manipulative state intervention, seeking to shape citizens' choices unnoticeably (Goodwin 2012). Nudge – typically seeking to change citizens' behaviours through tapping into their unconscious, *System I* decision-making – is perceived as a threat to people's autonomy and freedom to make their own conscious choices. This worry about a deepened form of 'governance by stealth' (Mols et al. 2015), trying to manage citizens without them

being aware of it, has been issued widely within academia (*inter alia,* Hansen & Jespersen 2013; Baldwin 2014; Bovens 2008). It also came to the fore in the public media, where *Nudge* co-author Sunstein was designated as 'the most dangerous man in America' in right-wing political circles (Beck 27 September 2010), and where BIT UK's launch was initially met with headliners such as 'Nudge nudge, say no more. Brits' minds will be controlled without us knowing it' (Dunt 5 February 2014).

A related concern voiced by critics is that BI would imply the upsurge of technocracy. Hereunder fall worries about BI lacking transparency, public accountability, and democratic opportunities for participation and deliberation. Additionally, BI is said to be elitist. Starting with the latter concern, White (2013) notes in his *The Manipulation of Choice* that the BI movement tends to be overconfident in its assumptions about what citizens value and when they act 'irrational'. The Libertarian Paternalist approach, claiming to nudge citizens only towards choices in their own interests *as judged by themselves* (Sunstein & Thaler 2003), is implausible according to White. BI experts cannot know what really drives citizens in concrete choice situations and whether these citizens' choices are made rationally or not. These experts simply lack the local knowledge about these unique citizen lives. Hence, while BI's narrative is one of 'helping citizens in light of their own interests', the actual practice is one of a technocratic elite imposing its own values on its policy recipients. Furedi (2011) shares this concern. He views BI as part of a growing intolerance towards 'deviant' public behaviours. In his eyes, governments have not grasped to NBE for the truth value thereof but instead for its associated political and moral implications: that the help of 'enlightened' state elites would be needed to educate 'irrational' citizens and correct for their decision-making 'flaws' (Furedi 2011).

Following on from this, some critics have expressed worries about the denigrating type of citizen subject that tends to be promoted by BI. They claim that BI pushes an overly negative, inferior, demoted image of citizens that robs them of their human dignity (Whitehead et al. 2017). Campbell (2017) for instance argues that Nudge stimulates a discourse and policy practice in which citizens are harmfully treated as 'mugs'. And Furedi talks about the increasing 'infantilization' of citizens as the result of governments strategically interpreting NBE insights.

An additional technocracy-related concern is that BI policy, especially with regards to its 'non-educative' nudges that specifically target *System I* (Sunstein 2017a), is relatively difficult to observe for the public. These public choice architectures can affect human conduct 'in the dark' (Bovens 2008), and their underlying policy motives and normative assumptions can easily remain covert and implicit. As a result, it becomes difficult for citizens to democratically debate, contest and/or resist BI policy (Whitehead et al. 2017). As Mettler (2011) sees it, BI adds to a wider technocratic trend

of 'the submerged state', whose actions and motivations are increasingly hidden from citizens. This increasing invisibility makes it difficult to hold the state accountable. The problem, in other words, is that BI lacks opportunities for democratic control. Still in the process of experimenting and institutionalizing, BI practices have thus far been little 'disturbed' by legal bounds and accountability demands. Working in an ad hoc fashion and with high 'administrative discretion' (Lepenies & Malecka 2018), experts have been implementing all sorts of nudges that unlike more established instruments have not been subjected to formalized policy procedures. Although BI policymaking trend is indeed sometimes publicly debated *in abstracto*, a lack of accountability and transparency still remains at the micro-level, in the very concrete choice situations in which citizens are being nudged unnoticeably (see Bovens 2008 on *type* versus *token* transparency of Nudge policy).

Hence, to critics, the rise of BI means a problematic shift away from deliberative democracy. Instead of furthering a 'Think' agenda, devoted to creating deliberative platforms and engaging citizens in decision-making processes, BI pushes governments to choose the 'Nudge' route, more interested in citizens' behaviours than in their values and visions (see John et al. 2011 on 'Nudge' vs. 'Think'). This superficial behaviour change approach has (apart from the aforementioned epistemological shortcomings) a democratic deficit. As Button (2018) asserts, it only speaks to citizens' capacity to be passively influenced, instead of appealing to their more active 'civic capacity'. McLaughlin (2016: 105) shares similar observations, noting how BI reflects an anti-democratic trend in which true citizen empowerment is lacking:

> *Nudge*'s popularity reflects a process of the degradation of the radical, emancipatory roots of empowerment. *Nudge* is premised on the assumption of human irrationality and of the need for professional, expert guidance to help the masses negotiate the travails of life. This guidance of human action towards pre-determined governmental goals also entails the danger that public debate over what is the best way for society to develop is bypassed to the detriment of the democratic process.

Moving away from concerns about technocracy, another concern holds that the BI movement, although preaching an apolitical story about increased effectiveness and efficiency, is actually a disguised form of neoliberal, NPM-oriented politics. Critics claim that BI serves as a 'smokescreen' for government inaction (Bonell et al. 2011a). It strategically obscures the social responsibilities of the public and commercial sector, while pushing a politically 'safe' policy strategy centred on the targeting of citizens with relatively cheap and unobtrusive nudges. As the critics view it, the focus on

individual behaviour change is not instrumentally but ideologically motivated, conforming to the hegemonic neoliberal belief in individual self-responsibility (Whitehead et al. 2017; Leggett 2014). Accordingly, critics note the flaws in that ideological choice, arguing that the sole focus on individual behaviour change leads to an unfair distribution of social responsibility, overly holding citizens (versus commercial organizations and/or governmental institutions) accountable for public issues.

Related to the critique of BI as disguised neoliberal politics is a body of Foucauldian critiques of the behavioural state. These accounts employ a 'governmentality' perspective, developed by Michel Foucault (2007) in his 1977-79 *Security, Territory, Population* lectures, analysing how neoliberal states leverage NBE insights and associated techniques to increase their control over public behaviours. The argument here is that BI is in subtle ways pushing a neoliberal 'citizen subject'. Ossewaarde (2010) explains how BI discourse employs 'freedom of choice' as an instrument in creating an 'active', self-responsible citizen subject. Mulderigg (2017) comes to a similar conclusion in her critical discourse analysis of a Nudge-informed public health campaign. She evidences how this campaign subtly disciplined citizens into self-responsible subjects through the use of certain language, frames and symbols, thereby further marginalizing 'problematic' low-income groups while obscuring responsibility of corporate actors. Han (2017) also analyses how the discourse of individual choice reflects a new policy style that exercises power *through* freedom. He relates BI to a new type of neoliberal 'psychopolitics' that understands that moulding citizens into active subjects is far more powerful than trying to steer them by means of external state incentives.

*The way forward*

In short, critics of BI have on the one hand voiced epistemological criticisms, in which they criticize BI for being too shallow and not powerful enough, only able to change behaviours at a superficial level (e.g. Mols et al. 2015; Rowson 2011). Interestingly, another group of critics have shared moral criticisms that go in a different direction. Informed by governmentality theory, they assert that BI actually is (or might become) *too* powerful. It poses a threat to the autonomy of citizens, revitalizes an elitist technocratic model of policymaking, and subtly shapes human subjectivity at a discursive level, pushing an overly active and self-responsible citizen subject.

For the critics, the way forward is to dismantle the behavioural state or at least prevent it from growing further. Some critics have also plead for a transformation of the behavioural state. In *Transforming Behaviour Change,* Rowson (2011) calls for a more participative and complexity-minded type of behavioural state practice. He notes that we need to move from 'Nudge', which treats complex public problems as mere

managerial issues, to 'Steer', which treats these problems as adaptive issues that require public deliberation. Similarly, John (2018) has postulated the agenda of 'Nudge Plus', urging us to rethink the cognitive and moral foundations of the behavioural state. In John's view, 'Nudge' doesn't necessarily have to be executed in a technocratic fashion, with government elites leveraging behavioural science to control citizens. It is possible to make the behavioural policy process, at various policy stages, more participatory. Jones et al. (2013) have put forward another transformation agenda, called 'Punch!', calling for a behavioural state that is more epistemologically open-minded as well as more morally sensitive to its impacts on the changing power dynamics between the state, corporate and citizen domain.

***Sceptics***

Although advocates and critics appraise the emerging behavioural state very differently, they do agree about one thing: that this trend is substantive, 'here to stay', and worthy of debate. It is precisely this attached importance that is denied by a third, sceptical trench in the debate. This trench questions whether there actually is a behavioural state, and to the extent that it is there, whether it is really novel, substantial and influential in the long run. Sceptics are little impressed by grandiose promises of a so-called 'behavioural turn'. Rather, they see such claims as part of a superficial Nudge hype: the 'latest cult US import' (Fitzpatrick 2011) blowing over. There is some common ground between sceptical and critics in the sense that they are both dismissive of the behavioural state. However, they are so for substantially different reasons. The critics ground their dismissal in epistemological and moral considerations. The sceptics, instead, are concerned with BI's impact/permanence, novelty and conceptualization – or rather, the lack thereof. These sceptical concerns are clarified below.

*Scepticism about impact*

To begin with, sceptics claim that the policy impact of BI is vastly overestimated. They argue that BI, while presented as a rapidly globalizing movement with a new grand narrative, by and large still is a marginal phenomenon (Whitehead et al. 2017). According to Selinger and Whyte (2012), BI's policy rhetoric 'was always inflated' and has not generated any real policy impact. Schlag too (2010) states that Nudge has actually had little political significance. It merely concerns the optimization of already made policy choices. For instance, Nudge tells us to readjust school canteens but it doesn't think more fundamentally about who should bear the responsibility for healthy school canteens. In a similar fashion, Selinger and Whyte (2012) argue that BI has thus far mostly been about the use of nudges as 'techno-fixes' of which the policy value is inevitably limited. BI's lack of impact, as sceptics argue, is also due to its limited scope of application, featuring only a select number of possible policy applications that seem to be copied over and over again in the field. Even BIT UK, the

field's frontrunner, as Ryan (2017: 3) observes, 'seems to have exhausted its policy ideas'.

*Scepticism about novelty*

Sceptics also challenge the novelty of governments' behavioural turn. They assert that many of the newly introduced ideas, such as 'nudges', 'choice architecture' and 'Libertarian Paternalism', are in fact 'old wine in new casks' (e.g. Grune-Yanoff 2012; Yeung 2012). In the sceptics' eyes, the dominant story of how the behavioural state emerged out of the popularization of NBE insights within policy spheres since the 2000s, fails to do justice to the more complex history of the behavioural state. This history has scientific and institutional origins that go back much further in time. Let's first consider BI's scientific legacy. Sceptics argue that the theoretical ideas by which BI is inspired (i.e. NBE) are hardly innovative. Both governments and academic institutions have been gathering knowledge about human behaviour, including its unconscious drivers, long before NBE flourished and became mainstream in the policy realm. Broadly speaking, a behavioural orientation has long been a core characteristic of many scientific disciplines, including economics, sociology, anthropology, neurology, biology, psychology, communication studies, and design sciences. But also in a narrower sense, various schools of thought have developed highly similar ideas to those later generated by the NBE school. Oliver (2013: 686) traces BI's legacy back as far as to the 18th century with Adam Smith's *The Theory of Moral Sentiments*, which like Kahneman's *System I* and *II* theory also distinguished between a 'rational "impartial spectator" and human actions that are often driven by particular affects.' In more recent history, there has been the psychoanalytic school of thought with its theorizing on the role of the unconscious in human behaviour (Gellner 1985). Additionally, as Rayner and Lang (2011) note, the idea that social norms shape human conduct had already been articulated by social scientists as early as in the 1930s, which was then taken up by the US advertising business. There also is the work of Janis and Mann (1977) on the cognitive, motivational and social mechanisms that humans use to cope with the burdens of choice. The most obvious precedent of NBE however has been the work of Herbert Simon (1945; 1957; 1985), whose ideas on *bounded rationality*, *satisficing* and the *homo psychologicus* laid the foundations of the *old school of behavioural economics* (Sent 2004). Simon, like the NBE scholars after him, was also interested in the cognitive decision-making patterns underpinning human behaviours and he also studied how humans make decisions in a context of limited time, information and cognitive capacity. He was awarded a Nobel Memorial Prize in Economic Sciences for his foundational work in 1978, suggesting that the ideas emerging from NBE were not only substantially speaking hardly new, but also that these ideas had already attracted significant popularity and public visibility.

Besides questioning the novelty of BI's scientific content, sceptics also argue that there is hardly anything new about BI's associated institutional practices. To begin with, they claim an orientation on behaviour and behaviour change comes natural to governments. It is in their nature to achieve policy goals and thus to aspire behaviour change. All government implicitly *is* behavioural government. But also at a more explicit level, sceptics note how typical elements of BI's institutional practice – e.g. field experiments, nudges, target group analyses – have long been used by governments. The history of modern public administration includes countless illustrations of behavioural policy instruments, programs and communities. Schuppert (2016) for instance notes how long before Nudge discourse gained traction numerous related techniques – e.g. persuasion strategies and ruling through signals – were already being used in government. He concludes that 'the concept of nudging cannot rightfully claim to have any news value'. Similarly, Esmark (2019) argues that BI's appropriated nudge-techniques are *de facto* no different from well-established communicative tools, including notification, moral suasion, persuasion, exhortation and public campaigns. Although he views such techniques in and of themselves as valuable, he is sceptic about the BI movement itself and its misleading 'branding savvy' (Esmark 2019). In the area of environmental policy, Graf (2019) shows in his 'Nudging Before the Nudge' paper how behavioural policies were already developed by the German government of the late twentieth century. Lastly, in the domain of public health, Bonell et al. (2011b) state that this domain already has a rich history with behaviour change techniques, like motivational interviewing, peer education and structural/physical interventions. In their view, as BI doesn't add anything new in this respect, it should not set off a whole new research agenda:

> we shouldn't rush into investigating the evidence base of nudging unless it offers something that existing approaches do not. Defined negatively, nudges seem to be anything other than just giving people basic factual information to enable them to make more rational, conscious decisions, or compelling them to change behaviour. It isn't clear how nudges are distinctive in any other way. Public health is rarely coercive (other than to prevent harm to third parties), generally goes beyond information giving, and already seeks to influence how choices are presented. (Bonell et al. 2011b: 241-242)

In addition to a well-established use of behavioural policy techniques and programs, there also exists a longer history of behavioural policymaking communities, like BI. These communities were connected to particular, at the time popular strands of behavioural science. For instance, the field of Public Relations, emerging in the early 20th century and championed by the work of Edward Louis Bernays (1928), drew on psychoanalytical theory to control mass behaviour. The field of Social Marketing,

originating in the 1950s but really picked up on since the 1980s, has been has been another major behavioural public policy wave (Pykett et al. 2014). Since the 1990s, the EBP agenda also brought state-of-the-art behavioural science methodologies into the policy realm (Cabinet Office 1999). Within this wider history, sceptics argue, the present-day BI movement – with its 'Behavioural Insights Teams', 'Behavioural Science Officers' and 'Choice Architects' – is nothing more than simply the latest wave of behavioural public policy.

*Scepticism about conceptual clarity*

Besides scepticism about scoping and novelty, sceptics also have problems with how BI is conceptualized – or rather, how it lacks proper conceptualization. Particularly problematic in their view is the notion of Nudge (e.g. Oliver 2015; Gigerenzer 2015; Bonell et al. 2011b; Hausman & Welch 2010; Ryan 2017). Sceptics note that this is a rather broad concept that has been defined very open-endedly, i.e. as 'any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives' (Thaler & Sunstein 2008: 6). The precise instrumental nature of Nudge remains unspecified in this definition, which has sparked unproductive debates about what is and is not a nudge (see Jellema et al. 2014 on this 'Nudge Fuzz'). It has also led to government practices 'falsely' claiming to be nudging. Ryan (2017) for instance observes how even BIT UK, the field's leading role model, has disapplied Nudge theory. Ironically, also several of the nudge examples mentioned in *Nudge* don't match the book's own nudge definition, as noted by Hausman and Welch (2010). Oliver (2015) too voices scepticism about BI's conceptual clarity (although he is still in favour of using NBE in policy). He is particularly sceptic about the concepts of Nudge and Libertarian Paternalism, claiming that these overly evangelized yet shallowly adopted ideas have become fuzzy and non-directive. They muddle the debate and stand in the way of a clearer debate in which rhetorical concepts actually correspond with the world of practice:

> [T]here ought to be clarity and consistency in and between what one is proposing and what one is doing. Many purported nudge interventions in UK policy do not seem to comply with the original tenets of libertarian paternalism. In particular, they often seem to be motivated by externality concerns, and they sometimes appear to be informed by rational choice theory. Whether or not one agrees with the approach adopted by governments, they ought to be held accountable for their rhetoric. (…) the nudge label is being oversold by many researchers, and overbought by some in policy, such that the intellectual clarity of the approach has been lost. (Oliver 2015: 713)

Besides Nudge and Libertarian Paternalism, also the concept of 'behavioural insights' is problematized by sceptics. This notion leaves unclear which exact scientific (sub)disciplines are included in this body of knowledge. Generally it is acknowledged that NBE forms BI's core knowledge foundation. However, also insights from social and cognitive psychology, neuro-economics and design sciences, and even more broadly, anthropology, biology and other domains are sometimes included (e.g. Whitehead et al. 2014). This absence or at most implicitness of disciplinary demarcations in BI discourse raises sceptical questions about what exactly is the epistemic content of this field and how it is distinctive (for more on these 'boundary issues' see Chapter 6).

*The way forward*

Put briefly, sceptics challenge whether BI really is impactful, novel and conceptually coherent. They point to the movement's still marginal nature, and its 'magic concepts' which although rhetorically strong ultimately lack a clear and practical substance. In addition, sceptics argue that BI's institutional practices are hardly new and that its ideas are mere fancy reformulations of already available knowledge. In their eyes, BI suffers from a strategic collective amnesia, searching to legitimize itself through false claims of novelty.

For the sceptics, the way forward should be to deflate the inflated 'behavioural turn', and expose it as the latest fad. An analogy can be made with the folktale of 'The Emperor's New Clothes', in which an emperor was proudly showing off to his people what he believed was a new, although for incompetent people invisible, suit of clothes. While his followers, desperately not wanting to be seen as incompetent, all applauded him, one child eventually shouted out that the emperor was in fact not wearing anything at all. Like this child, the sceptics' desire is to reveal the 'nakedness' of the behavioural state. Once its rhetorical nature has been exposed, scholars and policymakers can concentrate on actually innovative ideas again. Also, uncovering the 'pre-history of nudge' (Vallgarda 2012) will enable scholars and policymakers to capitalize more comprehensively on the vast body of behavioural knowledge that has already been accumulated over time.

### *Mapping the trenches*

Table 2.2 presents an overview of the different trenches in the behavioural state debate.

**Table 2.2: Trench warfare about the behavioural state**

| *Trench* | *Advocates* | *Critics* | *Sceptics* |
|---|---|---|---|
| *Appraisal* | *Epistemological advocacy*<br>- BI policy is more effective and efficient<br>- BI policy is more realistically theoretically underpinned<br>- BI policy is based on higher levels of methodological rigour<br><br>*Moral advocacy*<br>- BI forms an attractive middle ground between laissez-faire liberalism and coercive paternalism<br>- BI has a more realistic appreciation of the value of individual choice<br>- BI policy is deemed acceptable by a majority of people | *Epistemological criticism*<br>- BI policy is non-, semi- or counter-effective, or unevidenced in the first place<br>- BI's methodology is narrow and shallow<br>- BI is naively rationalist about the role of knowledge in policymaking<br><br>*Moral criticism*<br>- BI policy threatens autonomy<br>- BI is technocratic<br>- BI pushes a denigrating view of the citizen subject<br>- BI is disguised neoliberal politics, subtly disciplining citizens into active subjects | - BI lacks size, impact, permanence, and application possibilities<br>- BI is not novel but mere new rhetoric for existing practices<br>- BI is fuzzy and underconceptualized |
| *The way forward* | Fulfil the promise | Mitigate the pressure | Expose the phantom |
| *Key players and works* | *Inter alia,* World Bank 2015; OECD 2017; Dolan et al. 2010; Service et al. 2014; Sanders et al. 2018; Halpern 2015; Van Bavel et al. 2013; Lourenço et al. 2016; Lunn 2014; Shankar & Foster 2017; John 2018; Thaler & Sunstein 2008; Grimmelikhuijsen et al. 2017; Allemano & Sibony 2015; Shafir 2012. | *Inter alia,* Rowson 2011; Jones et al. 2013; Shove 2010; Whitehead et al. 2018; Lodge & Wegrich 2016; Goodwin 2012; White 2013; St. Paul 2011; Furedi 2011; Campbell 2017; Mettler 2011; McLaughlin 2016; Han 2017; Button 2018; Mulderrig 2018. | *Inter alia,* Ryan 2017; Fitzpatrick 2011; Selinger & Whyte 2012; Schlag 2010; Kosters & Van der Heijden 2015; Vallgarda 2012; Rayner & Lang 2011; Schuppert 2016; Hausman & Welch 2010. |

***Out of the trenches onto the field***

This chapter has sketched the background and reviewed the scholarly debate with regards to the present-day behavioural state. It has pinpointed the puzzling diversity of viewpoints in the academic trench warfare about this state. Advocates portray it as

a great promise for a finally effective yet not too obtrusive government, while critics either view it as methodologically limited and ineffectual, or instead too powerful – subtly pressuring citizens into behaving and understanding themselves as active subjects. Sceptics reject the premise of a 'rising behavioural state' altogether, explaining this trend as at most a marginal, faddish, fuzzy phenomenon. This trench warfare brings us at an impasse: the behavioural state can hardly be a promise, pressure and phantom at the same time. To move beyond this impasse, empirical light needs to be shed on what its early adopters are actually doing. The time has come to move out of the trenches onto the field. So, let's examine what really goes on inside the behavioural state.

# II | Observations

*Chapter 3*

# Brokering the Central Behavioural State

*A behavioural insights community has emerged within a growing number of governments. While this community helps to make policies more behavioural science based, its frontstage role models tend to assume a straightforward, instrumental and apolitical view of the science–policy relationship that seems unrealistic. This chapter therefore examines what goes on backstage in this community, based on an ethnographic study of behaviour experts in Dutch central government. The chapter argues that their work consists of a complex palette of practices (that is, choice architecture; analysis; capacity building). Because these practices resemble typical knowledge brokerage work, the chapter pushes for an envisaging of 'behaviour experts as knowledge brokers'.*

---

Chapter 3 is published as an original research article by Joram Feitsma in Policy & Politics.

Feitsma, J.N.P. (2019). Brokering behaviour change: the work of behavioural insights experts in government. *Policy & Politics, 47*(1), 37-56. doi: 10.1332/030557318X15174915040678

## 3.1 Introduction

Behavioural science, and in particular behavioural economics, has recently been seen as a promising source for better policymaking (Lunn 2012). This 'behavioural turn' manifests itself in governments' widespread consultation of behavioural economics bestsellers such as *Nudge* (Thaler and Sunstein 2008), *Predictably Irrational* (Ariely 2008), and *Thinking, Fast and Slow* (Kahneman 2011), as well as the appointment of leading behavioural scholars in key government posts, for instance *Nudge* co-author Thaler as strategic advisor to the United Kingdom's Cabinet Office. Most pivotally, though, the behavioural turn is visible in the trend towards specialized 'behavioural insights teams' (BITs) in Anglo-Saxon governments. These BITs form the 'frontstage' of behavioural policymaking and act as role models for how to translate behavioural science to policy. As special behavioural units, they present themselves as a new and exclusive policy profession, embodying its own knowledge, skills and identity. Central to this identity is the use of randomized controlled trial (RCT) methodology on subtle, psychologically-informed policy changes in order to optimise policies in an evidence-based way (John 2014; Service et al. 2012).

While acknowledging that the behavioural policy frontstage has been highly successful in putting behavioural science on the policymaking agenda, it also has a more problematic side. That is, it both implicitly and explicitly makes a number of problematic assumptions about the nature of the policy process and the role of evidence. More specifically, it tends to view the process of embedding behavioural science evidence into existing policy procedures as a rather simplistic matter of 'applying behavioural insights' (for example, Van Bavel et al. 2013; World Bank 2015; Hallsworth et al. 2016). This rather straightforward, instrumental, and apolitical take on the science–policy relationship has been widely critiqued (for example, Lindblom 1959; Simon 1985). As such, one may wonder how representative this behavioural policy frontstage is for what actually goes on backstage when policy actors try to feed behavioural science into their organization. In light of this question, this chapter aims to make two contributions, one empirical and the other theoretical. The empirical contribution is made by going backstage and generating 'thick' descriptions of the people inside the behavioural insights world beyond its spectacular and well-known frontstage. The chapter zooms in on an underexplored locus, Dutch government, where a behavioural insights community is slowly emerging within ministry buildings, but also beyond the deep state; from secondary school canteens to military training camps. By shedding light on actual, day-to-day practices, the chapter will show that the Dutch behaviour experts are misrepresented by the behavioural policy frontstage with respect to the complexity of their endeavours. From that observation, a theoretical contribution is made by linking the complexity of everyday behavioural

policy practice to the literature on knowledge brokerage (Hoppe 2010; Knight and Lyall 2013; Ward et al. 2009). Knowledge brokerage refers to the idea that before evidence is actually useable in policy settings, it first needs to be properly brought in, moved around, translated and contextualized. This chapter will argue that behaviour experts are better understood as such knowledge brokerage agents than as the direct choice architects envisioned in the literature.

The chapter first discusses the emergence of 'behavioural insights' in the global policy arena, and then briefly summarises theoretical debates on the use of evidence in policymaking. After elaborating on the ethnographic research approach, and revealing the emerging Dutch landscape of behavioural expertise, the chapter then highlights three key practices of behaviour experts: (1) choice architecture, in various forms, (2) analysis and (3) capacity building. The case is made that, in light of these particular practices, behaviour experts assume the role of knowledge brokers.

## 3.2 The long winding road from behavioural insight to policy

The interest of governments in behavioural science is hardly a novel phenomenon, but something that has appeared and re-appeared in many different guises. There are longstanding policy traditions that capitalise on theoretical and methodological insights from behavioural science, for instance drawing from Simon's (1985) account of the rationally bounded *homo psychologicus*, Tversky and Kahneman's (1974) heuristics and biases programme, social marketing theories (Pykett et al. 2014), behavioural policy design theories (Schneider & Ingram 1990), and evidence-based policy thinking (Cabinet Office 1999). Since *Nudge* (Thaler & Sunstein 2008), governments worldwide have launched various behavioural insights-related initiatives to enrich their policies with findings and methods from behavioural science (OECD 2017). Prominent behavioural economists became influential policy advisors, strategies to integrate behavioural insights into the policy process were formulated, and, inspired by the original BIT in the United Kingdom, special BITs were formed inside many governments, including Denmark, France, Germany, Singapore and the Netherlands (Lourenço et al. 2016). Furthermore, according to Whitehead et al (2014), two thirds of the countries worldwide have behaviourally-informed policies in one way or another, suggesting that this development is already widespread.

The behavioural policymaking trend has received substantial academic attention in the last decade (John 2014; Whitehead et al. 2017; Lodge & Wegrich 2016; Strassheim et al. 2015). Thus far, major themes in the study of behavioural policies include their effects and working mechanisms (e.g. Thaler & Sunstein 2008), legal implications (e.g. Alemanno & Sibony 2015), political meanings (e.g. Leggett 2014), and ethical

desiderata (e.g. Bovens 2008). A mainstream discourse has developed about what behavioural insights are, and how they should be put into policy practice. Core foci in this discourse are the discovery of nudge-interventions as a novel toolkit that recognises the more-than rational aspects of human behaviour, and the importance of *ex ante* evaluating of 'what works' with the help of RCTs. As such, the advocacy of behavioural policymaking goes hand-in-hand with that of earlier advocacy of evidence-based policymaking (e.g. Cabinet Office 1999).

It is important to recognise what the leading behavioural policy discourse implicitly assumes about the relationship between science and policy. That is, it tends to depict the integration of behavioural insights into policy practice as a straightforward, simple, and liberating act. It is not coincidental that in general the field talks about 'applying' behavioural insights to policies, exemplified by titles as *Behavioural insights applied to policy* (Lourenço et al. 2016), and emphasising the *simplicity* of that application, illustrated by titles such as *EAST: Four simple ways to apply behavioural insights* (Service et al. 2014). Such jargon points to a particular conception of behavioural policymaking as involving the simple, direct and rational-minded 'transferring' of evidence from science to the policy realm. The success of *Nudge*, for example, has come in part from its ability to make simple translations of academic behavioural insights to concrete interventions, resulting in a book that is not only littered with examples of successful nudges but also parsimoniously presents the basic 'principles of good choice architecture' with the help of the acronym 'NUDGES' (*in*centives, *u*nderstand mappings, *d*efaults, *g*ive feedback, *e*xpect error and *s*tructure complex choices) (Thaler & Sunstein 2008). While acknowledging that many reports from the behavioural insights frontstage do indeed make initial note of the complex processes behind behavioural policymaking, they nevertheless predominantly focus on extracting lists of simple 'principles' and 'tools' from the body of behavioural scientific knowledge rather than on going into depth about these complexities. For instance, a typical behavioural insights report in the context of health states that '[i]t seeks to arm the professional or policymaker with a simple set of tools that can be used to help shape patient or population health behavior for the better' (Hallsworth et al. 2016: 3). This instrument-oriented emphasis on application principles, examples and tools, implies that the appropriate role of behaviour experts would be that of *choice architect*, tweaking environments here and there in subtle ways. Moreover, they would do this while continuously testing 'what works', which points to another set of implicit assumptions, namely that it is both possible and desirable to produce fixed causal knowledge about the effects of policy changes, and that RCTs have the highest epistemic authority in doing so. The prevailing behavioural policy discourse thus makes at least three assumptions: the instrumental relationship between science and

policy; the stable and universal character of knowledge; and the hegemony of RCTs as the golden standard of evidence (Rouw 2011).

The abovementioned assumptions stand at odds with the literatures on policy translation (Ingold & Monaghan 2016), knowledge brokerage (Hoppe 2010; Knight & Lyall 2013), and evidence-based policy critique (Cairney, 2017). These literatures cast doubt upon a rationalist, instrumental and apolitical application of science to policy. To begin with, the idea that evidence will naturally find its way into the right policy actor's hands at the right time is contestable. Translating and circulating evidence requires the extensive work of connecting to relevant policy actors, speaking in their language, and meeting their needs. As contemporary policymaking is best understood as a disjointed process in which many actors – public and private, political and administrative – participate, with the governments in an increasingly distanced and *meta-governing* role (Sørensen & Torfing, 2009), the work of translating evidence has at once become more crucial and complex.

Additionally, it is important to recognise the role that bounded rationality plays within that disjointed policy process. Lindblom (1959) has demonstrated that policymakers are generally prone to stick to the status quo. Driven by the political imperative to act at the right time, they tend to satisfice with partial analyses, consider only incremental changes, and refuse to look at new types of evidence and methods. Moreover, as behavioural insights embody a big family of ideas (including controversies and conflicts), it can hardly be expected that policymakers incorporate these insights on their own, especially given that many of them do not have a background in behavioural science. Here too, translation work would be needed that surpasses a simple 'applying behavioural insights' mentality.

Furthermore, one can question the behavioural policy discourse's assumptions about the nature and hierarchy of evidence. That is, its elevation of causal 'what works' knowledge, and its assumption that such knowledge is fixed and universal, is problematic (Cairney 2017). A fundamental difference between science and policy is that while policy is future-looking and about 'what to do next' (Rip 2000), science is past-oriented (Kuhn 1962) and fundamentally uncertain about the future. It would therefore be more accurate to replace the evidence-based mantra of 'what works' with 'what worked' (Biesta 2007). Moreover, beyond this limited 'what worked' knowledge gained through experiments, there are various other, softer ways of gathering evidence that may provide policymakers with knowledges that experiments cannot produce. For instance, while experimental knowledge may show the behavioural effects of a certain intervention on a certain group at a certain place and time, qualitatively produced knowledge has the capacity to produce rich and

situated accounts of the underlying experiences, thoughts and mental life behind human behaviours that can help to explain such effects. Another important type of evidence that is little recognized by the behavioural policy discourse is local 'how-to' knowledge, which is a crucial requirement for policymakers to actually integrate generic knowledge in particular contexts (Rouw 2011). In this sense, evidence-based policies can at the same time be 'evidence-blind' to the extent that they exclude valuable sources of evidence other than RCT-evidence. Policymakers may benefit from incorporating a broader palette of evidence bases.

Last, the behavioural policy discourse seems to overlook the fact that the science–policy relationship is inherently politicized (Hoppe 2010; Lindblom 1959). It disregards that evidence is not necessarily brought in for the instrumental purpose of more rational policymaking, but also serves political agendas and interests that can easily overrule an instrumental use of evidence. More specifically, policymakers employ several techniques to deal with evidence, as they may cherry-pick the evidence they need and disregard the rest ('fish'); seek to exert influence over researchers to 'fabricate' desired evidence ('farm'); discredit evidence that is detrimental to chosen policy directions ('flak'); or place constraints on those actors who produce or promote such evidence ('strain') (Ingold & Monaghan 2016). Behavioural insights are not excluded from such politicization of evidence, but are also part of a process of being framed, moulded or neglected in order to satisfy particular political powers and interests. They are subject to the interplay between the fundamentally conflicting languages, rhythms and logics of science and policy. In introducing these insights into the policy system, one would not come far with a simple 'knowledge transfer' mindset. Here too, translation work is needed, mediating between the conflicting worlds of science and policy.

In light of the above-mentioned critiques, it becomes less plausible to view behaviour experts as choice architects who directly apply behavioural science. Instead, this chapter will argue that, based on the ethnographic findings of this study, behaviour experts are better understood as *knowledge brokers* (Meyer 2010). The notion of knowledge brokerage has arisen against the background of an increased acknowledgement of the complexity and conflict in the science–policy relationship, resulting in 'evidence-policy gaps' in which policies are formed without being attuned to the available body of relevant evidence. To diminish these gaps, a professional group of 'dedicated knowledge brokers' has made its appearance within governments, with official role titles like 'diffusion fellows', 'knowledge transfer associates', and 'chief science officers' (Kislov et al. 2016). Knowledge brokers can be understood as 'boundary arrangements' (Hoppe 2010), dedicated to the collection, diffusion, and translation of evidence so as to smooth the flow of information between

science and policy (Knight & Lyall 2013). The nature of knowledge brokerage can be described more specifically by the knowledge broker's three main tasks: information management (*gathering* and *transferring*); linkage and exchange (*networking*); and facilitation in turning situated knowledge into action (*transforming* and *facilitating*) (Kislov et al. 2016; Ward et al. 2009). This chapter will point out that the Dutch behaviour experts in central government also belong to this group of knowledge brokers – with the specification that they are *internal* knowledge brokers working with *behavioural* scientific evidence. In both their thinking and practice, they go well beyond a simplistic 'knowledge utilization' model. They recognise the need for extensive brokerage work, see themselves as the appropriate actors to meet that need, and organise themselves (for example, as special units between boundaries) and act (for example, training, networking and building tools) accordingly to this understanding of their role in government. Hence, the chapter argues for a renewed, more realistic conception of behaviour experts: as knowledge brokers instead of direct choice architects. Table 3.1 summarizes this contraposition and the underlying assumptions in it.

**Table 3.1: Behaviour experts as choice architects versus knowledge brokers**

| *Behaviour experts as choice architects* | *Behaviour experts as knowledge brokers* |
|---|---|
| • The science-policy relationship is straightforward and instrumental.<br>• Policymakers have unbounded rationality, time, and resources.<br>• Causal 'what works' knowledge is determinate and the golden standard of evidence.<br>• Policies are made by small, centre-staged policy clusters in which behaviour experts have relatively much controlling power over the choice architecture.<br>• Behavioural insights are easily converted to concrete applications (e.g. nudges) and do not require brokerage. Behaviour experts can focus on actual application (i.e. choice architecture) straight away. | • The science-policy relationship is chaotic, circuitous, and politicized.<br>• Policymakers have bounded rationality, time, and resources.<br>• Causal 'what works' knowledge is provisional, and there is a need for local, experiential, tacit 'how-to' knowledge<br>• Policies come about in wide, disjointed parts of (meta)governance in which the controlling power over the choice architecture is widely dispersed across actors.<br>• Behavioural insights first must be translated in line with the rhythm, logic, and language of policymaking before they are usable. And even then, continuous promoting, networking, and translating work is needed for actual take-up. |

### 3.3 Methods

This chapter starts from the little-used (exceptions include Whitehead et al. 2017; John 2014; Feitsma 2016) viewpoint that to understand the world of behavioural policymaking, it is important to study its people. From them we can learn how behavioural insights are being used in practice. Moreover, to get a deeper sense of what actually goes on inside the behavioural insights field, beyond the official stories told at the frontstage, it is important to study these people more intensively, longitudinally, in their natural habitats. Studying them from up close helps to uncover backstage realities (Van Hulst 2008) and provide a sense of the 'everydayness' (for example, typical rituals, routines, discourse and so on) which adds 'thickness' to our understanding of the field. This chapter adopts such an ethnographic approach, and falls within a longer tradition of 'administrative ethnography' (Boll & Rhodes 2015; also see for example, Rhodes et al. 2007; Van Hulst 2008, and even Kaufman 1960). Following the ethnographic principle of 'being there' (Rhodes et al. 2007), I set out to study behaviour experts 'out there', to see who they are and what they actually do. Over the course of 16 months (November 2014 until March 2016), I immersed myself in the worlds of behaviour experts and examined their everyday work practices, including their typical tasks, techniques, routines, tools and language.

My methodological toolkit consisted of talking, observing and reading (Rhodes et al. 2007). I started with ten unstructured, preliminary interviews with scholars and practitioners in the field of behaviour change. Then I performed 24 semi-structured interviews with 35 behaviour experts in the Dutch government. The interviews were guided by sensitizing topics, addressing the interviewees' *professional background, work relations, goals and tasks, everyday practices, successes and challenges*. Alongside interviewing, I observed behaviour experts in and out of their offices on 17 different occasions. The observations were short, up to five hours, totalling around 55.5 hours. More specifically, ten observations consisted in attending (internal) educational and knowledge exchange related events, five in shadowing behaviour experts during work, and two in unstructured conversations with behaviour experts at their workplace.

This type of 'hit-and-run ethnography' (Rhodes et al. 2007) – going in and out of the field, making short visits to different local sites – allowed me to observe behaviour experts in varied contexts, and balance being time-efficient with acquiring a sufficient degree of texture, nuance and depth in my observations. At the same time, a limitation of this hit-and-run approach is that it allowed less space for a very deep immersion in which the behavioural policy backstage, including its more shadowy parts, could

be captured more comprehensively. Rather, I have captured and reconstructed some of it. Throughout the research process I also studied relevant documents (public reports, presentations, e-mails, memos and so on) as a third source of data.

The presented findings flow from the aggregate of all interview, document and observation data. The interviews were recorded, selectively transcribed and turned into field reports that followed the structure of the sensitizing topics. Field notes during observations were also translated into field reports. Analysis and text work consisted in continually interpreting, comparing and reconstructing the data, *inter alia*, looking out for salient themes, noting differences and similarities, clustering information into categories, and translating initial fieldnotes into more elaborate case illustrations.

The case selection process started out with mapping the presence of behaviour experts in Dutch government. While later in the mapping process I used the snowballing technique and kept a list of existing behavioural units which I verified with respondents, I started with exploratory desk research. I looked at previous research (in particular Dorren 2015) on Dutch behaviour experts, and performed Google-searches (in Dutch) for, *inter alia*, 'nudging', 'behavioural insights teams' and 'applying behavioural insights', in combination with the name of particular agencies. Broad and varied search terms were needed, as the jargon of behaviour experts tends to vary and is still evolving, even though their practices are similar. Based on these searches, I screened through various content, such as reports, websites and online work profiles. As this initially resulted in a broad and blurry set of many potential 'quasi-behaviour experts' that only used behavioural science incidentally or implicitly, I followed some stricter selection criteria. I only selected self-proclaimed behaviour experts in Dutch central government who were structurally and explicitly using behavioural insights. Conversely, those who didn't profile themselves explicitly as behavioural science appliers, or only used them in an ad hoc or retrospective fashion, were excluded. Also, I only selected behaviour experts who worked directly for government. Within these boundaries, I selected a wide range of behaviour experts across policy domains and agencies, and included all of the relatively large behavioural units.

Behavioural Insights studies have shown that the global landscape of behavioural insights is rather differentiated (Whitehead et al. 2017). Within this fragmented landscape, the Dutch central government forms a comparatively low-profile case as most of the literature focuses on the Anglo-Saxon forerunners in the field, such as BIT UK (e.g. John 2014; exceptions are Lourenço et al. 2016; OECD 2017). Nevertheless, the Dutch government accommodates various kinds of emerging behavioural

practices which are important to study in order to understand behavioural policymaking in all of its varieties. The representativeness of Dutch behavioural practice for its international counterparts should, however, not be overstated. Rather, studying the Dutch case may help to shed more light on the role of institutional context in shaping unique varieties of behavioural policymaking. Unlike the Anglo-Saxon cases, the Dutch government has not yet deeply institutionalized behavioural expertise and shows more signs of an expert- and consensus-based policy culture rather than an evidence-based policy culture (cf. Strassheim et al. 2015). Such contextual differences are likely to affect what kind of behavioural policy practices emerge.

### 3.4 What Dutch behaviour experts actually *do*

***The Dutch landscape of behavioural expertise***

The Dutch central government exhibits an explorative yet widespread interest in behavioural insights. Thaler and Sunstein's (2008) *Nudge* was followed up with a string of reports on behaviourally-informed policy from official advisory bodies for the government, peaking with a memorandum to Parliament on the use of behavioural insights in policymaking (Ministry of Economic Affairs 2014). The vast majority of interviewees set up shop somewhere between 2009 and 2015, triggered by the increasing popularity of behavioural insights at the time. During that period, 'behaviour change' was put on many policy agendas, behavioural scientists were hired and various behavioural projects, units and networks were set in motion. As these developments are relatively recent, most behavioural practices are not yet deeply institutionally embedded but instead are organized informally, from the bottom-up, with limited resources and limited connections to existing policy actors and routines.

Behaviour experts work in many places inside government. The 35 behaviour experts I interviewed represented 20 different agencies alone, and during my observations I have met behaviour experts from many other departments and organizations. Some work in the direct ministerial centre of government, while others are further removed from it and part of independent public agencies with regulatory, enforcement, knowledge distribution or implementation tasks. They are also involved in a wide range of policy areas, which can vary from food waste to tax compliance, to give just two examples. This diversity suggests that behavioural insights have already seeped into central government to a considerable degree.

Behaviour experts have diverging professional backgrounds. Half of the interviewees, 17 in total, have received actual academic training in the behavioural

sciences, for instance in behavioural economics, (social) psychology, behavioural finance, communication sciences and criminology. The other 18 interviewees' original professional backgrounds are less straightforwardly connected to their current behavioural practice, with a predominance of social scientific backgrounds (for example, political science, law and public administration) but also backgrounds in the humanities (for example, philosophy) or beta-sciences (for example, chemical technology). These non-behaviourally trained interviewees tend to rely on a more basic level of behavioural scientific knowledge, mostly acquired through postsecondary training and self-study, reading popular works like *Nudge* (Thaler & Sunstein 2008) and *Thinking, Fast and Slow* (Kahneman 2011).

There is also diversity in the organization of behavioural expertise. In some cases, there are specialized behavioural insights units that conform to the BIT UK model. I identified five BITs in Dutch central government, all comparatively small, with up to five members, and new, all being founded after 2008, with official names like 'BIT' and 'Team Behaviour Change', or informally calling themselves 'BITs'. Yet, the Dutch behavioural landscape includes other, smaller, more explorative kinds of specialized practices, such as knowledge networks (for example, an interdepartmental 'Behavioural Insights Network'), research programmes, and work groups. Also, several individual behavioural functions (for example, 'Behavioural Insights Advisors') have been installed. In other cases, no specialized 'behaviour expert' functions are put in place but behavioural insights are instead integrated into existing organizational processes.

Table 3.2 summarises the background of the interviewed behaviour experts and their organizations. It gives a non-comprehensive snapshot of the dynamic behavioural landscape in Dutch central government from early 2017.

**Table 3.2: Background characteristics of the interviewees**

| | |
|---|---|
| *Organizational setting* | 9 working in ministries<br>6 working in executive agencies<br>5 working in regulative agencies |
| *Organizational design* | 5 BITs<br>8 other exclusive designs<br>7 integrated designs |

| | |
|---|---|
| *Starting point* | 15 started since or after 2009<br>5 started before 2009 |
| *Professional background* | 17 in behavioural sciences<br>13 in social sciences<br>2 in humanities<br>3 in beta-sciences |

The chapter proceeds with an account of everyday behavioural policy practice. This account challenges the representation of behavioural expertise by its frontstage models. While behaviour experts indeed sometimes follow these models, namely when they act as direct and solo choice architects, overall their practices turn out to be more indirect, corresponding better with knowledge brokerage activities. The chapter highlights three of their key practices, which, although not mutually exclusive, reveal substantially diverging approaches: (1) choice architecture, in various forms, (2) analysis and (3) capacity building.

### *Choice architecture*

*Solo choice architecture*

One key practice of behaviour experts is that of carrying out all sorts of concrete behaviourally-informed interventions in people's physical and informational environments, also known as 'choice architecture' (Thaler & Sunstein, 2008). Occasionally this practice mimics the frontstage models, namely when behaviour experts make these kinds of minor environmental readjustments autonomously, acting as *solo choice architects*. For example, they might try to steer policy subjects by highlighting certain individual choices, reframing them, providing social feedback, and so forth. Looking at the Dutch behaviour experts, perhaps the case illustration below comes closest to living up to the ideal type of choice architecture. The illustration describes a 'school canteen officer', working at the Netherlands Nutrition Centre Foundation, who helps to make school canteens healthier. She is a choice architect *optima forma*, as she literally travels from choice architecture to choice architecture, canteen to canteen (notably also the first given example of choice architecture in *Nudge*), inspects them, and then suggests many small design-led changes to stimulate healthy behaviour.

> Inspector Nudge
> I meet up with the school canteen officer on the empty parking lot just in front of the school, where she is about to start her visit. As one of the eight fulltime members of the 'School Canteen Brigade', she visits schools

> throughout the whole country to assess how healthy their school canteens are and consult school directors and canteen managers about what could be changed to help pupils make healthier food choices. She has been doing this work for three years now and has made over 200 visits. This morning she is visiting a small secondary school in Apeldoorn. After she receives a short tour through the school building from the canteen manager, we arrive in the, at that time, empty school canteen. Only three schoolboys are standing behind the canteen counter, preparing sandwiches for the coming lunch break. The canteen officer looks around for a while. She takes out her iPad and starts taking pictures of the canteen, the counter, the menu list with food prices, the plate of sandwiches, the bowl of fruit in front of it, and a tap water point next to the counter. We sit down at one of the tables. The canteen officer unfolds her map with folders, stuffed with information, tips and tricks to make school canteens healthier. Then she starts to give some of her observations to the manager. She suggests looking at various aspects, including the volume of healthy versus unhealthy products, the appearance of products, their availability and the power of peer influence (it matters when 'the coolest guy eats an apple'). Meanwhile, she mentions all sorts of little tricks: placing healthier products more prominently, making attractive offers for healthier products, presenting fruit in a nice fruit bowl or precutting it into smaller pieces, emphasising healthy products in the menu list, and numerous other nudge-like techniques. During the lunch break the canteen manager immediately follows up on one of her suggestions: to make water more available. He asks the schoolboys behind the counter to fill some jugs with tap water and place them, with some cups, in front of the counter and promote the free water among the pupils. Not long after that, the first pupils have poured themselves a cup. Behaviour change, apparently, can be that simple.

Solo choice architecture is also practised by those behaviour experts who autonomously run small-scale field experiments. In applying the classical behavioural scientific method of the RCT, they closely follow the forerunning BIT, which has made RCTs the trademark of its approach (Haynes et al. 2012). These RCT-oriented experts tend to have a strong background in economics or social psychology, and employ a modernist– empiricist scientific language (for example, talking about 'hypotheses' and 'treatment groups'). Their ambition is mainly to discover which policy interventions actually work. They believe that theoretical assumptions offer a poor basis for policymaking, as the behaviour of policy targets is too complex to predict beforehand. That is why these experts argue that new interventions should be

empirically tested, and field experiments are the best way to do so. One interviewee stated:

> It's more about the general vision that you'd want policymakers to have. That they don't start from assumptions made in advance, but that they have the courage, the guts to put into question what the behaviour would be like. And to find that out, you'll want to do experiments. That's where we're trying to pioneer a little.

Dutch behaviour experts have already run several trials on behaviourally-informed policy changes. For instance, the BIT at the Dutch Tax and Customs Administration, a typical follower of the RCT approach, managed to increase tax compliance with roughly 10 per cent in a field experiment, sending out letters to tax payers with minor changes appealing to Cialdini's (1984) social influence mechanisms (for example, scarcity, liking and reciprocity). However insightful and potentially effective RCTs can be, conducting them also tends to be difficult. They are highly technical processes that are time- and labour-intensive, costly and require a lot of coordination with relevant stakeholders in the field.

The hardships behind the RCT philosophy help to make sense of the observation that not all behaviour experts adhere to it. Some replace this costly, technical and therefore hardly feasible approach with a more *pragmatic* approach. They then develop behaviour change strategies based on 'educated guesses', grounded in field observations, existing scientific literature and common-sense reasoning. Their focus is not so much on the methods of behavioural science as it is on its theoretical contribution (for example, general knowledge of heuristics and biases). Instead of hard experimentally-tested evidence, they work with softer evidence and make 'estimations'. While the downside of this approach is that it negates part of the inductive, evidence-based spirit ingrained in behavioural science, a major advantage is that it allows for intervention at a much higher pace, with much more freedom. To illustrate: while interviewees only mentioned 11 distinct field experiments, many of which were unfinished, the amount of interventions that followed from a pragmatic approach was considerably higher. These interventions included default changes (for example, removing automatic maximum loaning options for student loaning), gamification (for example, turning the job-seeking process into a game with 'expedition work') and physical space readjustments (for example, designing for dialogue-stimulating layouts of company meetings). Most interventions were of an informational kind, embedded in letters, websites, mailings, conversations or text messages. To further illustrate the pragmatic approach: when I attended an introductory course in behavioural insights at one of the ministries, I met an intern at

a regulatory agency who instead of rigorously testing her nudges with RCTs, simply suggests minor nudges for colleagues to implement in ongoing projects. Inspiration comes from frontrunners like BIT and she freely uses their examples. She also has developed a 29-page-long list of behavioural techniques from which to draw upon. Some of these techniques are perceived as basic and universally applicable nudges. Outgoing letters from judicial departments are especially 'nudgeable', as they tend to be written in a foggy language that doesn't help to achieve the desired behaviour change. Thus, it seems that when approached pragmatically, solo choice architecture is a varied and adaptable policy toolkit.

*Co- and contra-choice architecture*

From the observations and interviews, it appears that there are at least two ways in which behaviour experts deviate from the above-mentioned solo choice architecture. First, rather than designing and carrying out interventions autonomously, they do so more often in joint effort with other parties. This collaborative nature is to some extent even the case for the canteen officer described earlier, since she doesn't actually redesign school canteens herself but rather consults those who do. Like her, many behaviour experts are at most *co*-choice architects, dependent on many others (including universities, consultants and colleagues) to get things done. Some interviewees even fully take on a project management role, solely handling matters of supervision and coordination while outsourcing practical and research-related tasks to others.

Second, behaviour experts do not always operate by designing their own choice architectures that directly affect policy targets. They also inspect and regulate the choice architectural designs of commercial businesses. This is particularly visible within the context of regulatory agencies, where behaviour experts investigate whether businesses are not using behavioural insights in ways that impair the decision-making of consumers. When these experts do indeed identify wrongful uses, they undertake actions to undo or reverse them. These kinds of 'counter nudging' (Alemanno & Sibony 2015) practices can, for instance, entail that behaviour experts put businesses under pressure – sometimes threatening with sanctions – to install decision horizons, give more honest or transparent information, and remove harmful anchors and defaults. Concrete examples are the pressing of businesses in the travel industry to have no pre-checked boxes installed that make consumers purchase additional travel products by default, or urging 'Booking.com' to give more honest information about the availability of hotel rooms (Whitehead et al. 2017). In regulating commercial choice architectures, behaviour experts become *contra*-choice architects, protecting consumers against behaviourally-informed harm.

Choice architecture is thus underpinned both by different underlying roles (that is, solo, co- and contra choice architecture) and different views on what is useful evidence. Table 3.3 recaps how these views result in two contrasting approaches.

**Table 3.3: Different approaches towards choice architecture**

| | *RCT-approach* | *Pragmatic approach* |
|---|---|---|
| *Focus* | *Ex ante* policy evaluation | Policy advice and ad hoc intervention |
| *Output* | Some small-scale experiments | Many small (permanent) tweaks |
| *Method* | Behaviourally-tested (RCTs) | Behaviourally-informed (field research, literature study, common sense and professional assessment) |

These different approaches closely align with the dichotomy of behaviour experts as choice architects versus knowledge brokers. The group of rigorous field experimenters who aim to make evidence-based policies clearly advocate a choice architecture perspective, which manifests itself in a rather exclusive appreciation of causal, 'what works' knowledge gained through RCTs. Also, to some extent their approach assumes unbounded time and resources, given the highly labour-intensive work that is required to run a single trial. However, there is also a group of behaviour experts who use behavioural insights more pragmatically: more 'behaviourally-informed' rather than 'behaviourally-tested' (Lourenço et al. 2016). Their approach resembles a knowledge brokerage perspective as it doesn't assume a hegemony of RCT-knowledge but instead draws on a plurality of softer sources of evidence, partly out of a recognition of limited time and resources. Both the co- and contra-choice architecture also tie well with this knowledge brokerage perspective because both roles seek to include, address and intervene on the wider fields of governance of which behaviour experts are a part. As such, co- and contra-choice architecture substitute the ideal type of the solo choice architect role for a more distanced and meta-governing role, focused on building networks and working *with* and *through* policy actors.

***Analysis***

The take-up of behavioural science in policymaking is often, as in the previous section, associated with the instrumentation and implementation stages. Yet this knowledge is also incorporated at earlier stages in the policymaking process. Policy formulation, for instance, is a stage that receives the attention of the studied behaviour experts as they believe that policymakers often design ill-informed policies based on incomplete analyses and misguided lines of reasoning. The choice of policy instruments can be particularly ad hoc. According to one interviewee, policymakers

'just do something', based on loose speculation, habits and gut feeling. To prevent such poorly informed policy design, a second key practice of behaviour experts is *analysis*. This practice seeks to unmask underlying assumptions about policy targets and helps to produce richer and more empirical underpinnings of policies.

As explained during interviews and in official documents, these analyses tend to be done in a structured way, following a series of steps. The analyses usually start with the selection of a complex policy case with a strong behavioural component. Behaviour experts then take time to get to the bottom of the underlying policy theory and all the behavioural factors that may play a role in the selected case. They ask fundamental questions that may look self-evident but are often neglected by policy designers. Usually this process starts with 'demarcating' the policy problem, precisely defining the problem, the target group, and the desired alternative behaviour. Then the analysis focuses on what drives policy targets to behave as they do: 'Who are they? What moves them? What drives them?' In each case, behaviour experts will search for the 'origins' of the problematic behaviour: the 'behaviour determining factors'. Dependent on the nature (for example, cognitive, motivational, environmental) of these determinants, behaviour experts accordingly assess which behavioural mechanisms, strategies and instruments are well-suited to steer policy targets in the preferred direction. Thus, behaviour analyses examine both the determinants of target behaviour as well as the potential mechanisms through which to change that behaviour. Analysis and intervention are tightly coupled:

> I think that we're mainly looking for problems or things that do not go well in our provision of services, which potentially have a behavioural component as the origin. So, we identify the origin…And then we start to think: is there something in there…that we can change or improve?

Sometimes, these analyses are done with the help of special tools. For instance, one respondent organises group sessions over two or more days to play a serious card game, called the 'Behaviour Test'. This game takes the group along the process of thoroughly analysing a selected policy problem. By playing different cards that consider the potential role of specific behavioural insights, the group gradually develops more insights into the underlying drivers of target behaviour and effective intervention strategies. Besides the 'Behaviour Test', behaviour experts use many other tools and games, such as BIT UK's (Service et al. 2014) 'EAST' model in the form of a deck with inspiration cards, or the 'Campaign Strategy Instrument', which provides step-by-step guidelines and worksheets to design behaviourally-aware public information campaigns. That these step-by-step analytical tools for groups are welcomed so readily by behaviour experts is not surprising since they follow the

experts' own philosophy: if you want people to do something, make it easy, attractive, social and timely.

Behaviour experts draw on different research methods in their attempts to carefully 'reconstruct the decision-making of actors'. Some methods are more inductive, for instance: communication in the field, surveys, interviews, focus groups and observation. Others tend to be more deductive, such as: expert brainstorm sessions, literature study and logical reasoning. The behaviour experts that work strictly deductively seem to adopt an engineering attitude towards the nature of human behaviour and the extent to which it can be crafted. They view behaviour change of policy targets as a mechanical matter of 'finding out what makes them tick' and then 'pushing the right buttons'. It requires proper diagnosis of what drives the behaviour of policy targets, 'knowing their triggers', and on the basis of that deciding what behavioural mechanisms to exploit: which 'buttons to push'. These 'buttons' for behaviour change tend to be drawn from condensed theoretical models, for instance 'the three buttons of neurologist Victor Lamme' (fear, social and greed, see Lamme 2016).

The key practice of analysis generally reflects the knowledge brokerage perspective on the uptake of behavioural expertise. First, the shift made from a focus on instrumentation and implementation to the earlier stage of policy formulation itself shows an awareness of the complexity of taking up new evidence in a policy-setting. This evidence is not as simply 'used' as a choice architecture perspective would assume. Instead, evidence is integrated into the policy process more comprehensively, not just in its end stages. Second, the way in which this integration is attempted, that is, through developing practical tools and guidelines, also reveals a recognition that behavioural insights are not straightforwardly applicable, but first must be set in line with the rhythm, logic and language of policymaking. Third, the consideration of a wide range of evidence sources during analyses (combining the inductive and deductive, the soft and hard, the commonsensical and academic) reflects a pluralistic view of evidence that also fits the knowledge brokerage perspective.

***Capacity building***

Behaviour experts are, as we have seen, not just direct choice architects who make small adjustments at micro-level, but they are also indirect choice architects who operate at the meso- and macro-levels. While analysis occurs at the meso-level, not actually intervening but nevertheless strategizing about possible interventions in particular cases, this section turns to a third key practice that occurs at the macro-level, much further removed from actual choice architecture: *capacity building*. This

practice involves behaviour experts acting as ambassadors for the broader behavioural turn within government. They question the self-evident nature of traditional policymaking, which they deem to be rather 'odd', 'bizarre' and even 'outright scary'. To attract allies in their 'battle for better policy', they make their colleagues and managers more behaviourally-aware through a range of knowledge dissemination projects: they talk to people in their network, give presentations, write booklets and organise research programmes. The most intensive missionary work of behaviour experts is done in their roles as trainers. Several interviewees have developed educational modules on behaviourally-informed policy which they offer to their colleagues. The case illustration below illustrates this trainer role. It features a commander of a research unit at the Royal Netherlands Army, giving an introductory lecture on 'behavioural influence' to a group of special forces.

> 'Weapons of influence'
> 8:00 am sharp. The commander begins. The atmosphere is one of relaxed attention. The setting is pretty ordinary: military officers sitting behind little desks, wearing their green-brown camouflage outfits, drinking coffee, looking at a PowerPoint presentation. The commander starts by explaining how our brain works and how it shapes our decision-making. He introduces Kahneman's distinction between System I and System II. Then, all of a sudden, he throws his water bottle towards a soldier on the second row. In a split-second the soldier raises his hands to protect his face. 'So, did you have to think long before trying to catch it? Now that's System I thinking!', the commander enthusiastically explains. His lecture is filled with little exercises and intuitive examples from marketing and behavioural science. The underlying mechanisms of behaviour change, as he explains it, are not hard to grasp. He makes them understandable with the help of parsimonious theories, like Cialdini's model of the six mechanisms of social influence, one of the commander's all-time favourites.
> 11:00 am. Time for the soldiers to apply the lessons learned in their own context: suppose that they had arrived at a local village, and wanted to establish a relationship with the local leader in order to acquire valuable information. How could they engage this leader in a behaviourally-informed manner? The militaries are given ten minutes to think, after which many ideas come up. They could start off with small-talk or wear a casual outfit in order to win the leader's sympathy. They could emphasise his role as the leader to make him feel important. They could communicate their own high status as 'the chief' to trigger his authority bias. As the commander argues, these ideas may seem trivial but they are crucially

> important because 'in the end, a military operation is a communication process'. In two months these special forces will depart on their mission to Mali. They will do more talking than fighting. It is thus not their skills in armed combat that they will need most but their behavioural savviness: their 'weapons of influence'.

While the behavioural policy frontstage tends to assume a direct and self-executing role of behaviour experts, the case illustration above shows a different picture. It demonstrates a behaviour expert who works more indirectly, not actually applying behavioural science himself but educating others in doing so. Several other interviewees follow this approach: they train, *inter alia*, regulatory inspectors, call centre operators, school boards and municipal officials to become better choice architects. Those actors who function as 'street-level choice architects' are especially important to reach, as they stand relatively close to the policy targets they influence and can exercise more direct control over local choice architectures.

The key practice of capacity building aligns closely with a knowledge brokerage perspective. This manifests itself, first, in the awareness of behaviour experts that they are only small parts of widely fragmented governance structures, in which the power to invoke behaviour change is dispersed across many different policy actors. As such, they recognise the need to connect and reach out to all these actors and ensure that behavioural insights are properly introduced to them. Behavioural insights are to be anchored in the whole policy system by working *through* these actors. Moreover, behaviour experts recognise that this requires extensive translation work. That is why they spend much of their efforts on writing booklets and developing courses, which do not merely summarise the academic body of behavioural insights but actually 'transform' (Meyer 2010) these insights into novel, 'brokered' knowledge that is attuned to policymakers' ways of thinking.

The vast majority of interviewed behaviour experts experienced most success in this area of capacity building. At the same time, it must be noted that these kinds of successes were also perceived to be easier to achieve. The fact that behaviour experts are able to elicit beginner's enthusiasm certainly doesn't mean that colleagues will actually follow up on the 'tough' methodology. An interviewee stated: 'It's super fun, until you have to do it for yourself.' A sole focus on capacity building may therefore not be sufficient for a government-wide embrace of behavioural insights. Behaviour experts may need to get closer to the action and take the lead where necessary. Therefore, a combination of different roles, taking the lead in complex projects but also encouraging others to work for themselves, might be the best shot at making the behavioural turn within government permanent.

## 3.5 Conclusion and discussion

While frontstage role models of behavioural insights lead the way in behavioural policymaking and help to further promote it, they may at the same time overlook its backstage complexities. That is, their underlying assumption that behavioural science can be straightforwardly transposed into a nudge-toolkit readily available to the policymaker, appears oversimplified. This chapter has, based on an ethnography of behaviour experts in Dutch central government, argued for a richer representation of behavioural policy practice. Table 3.4 captures this richness. What is important is that it shows that behaviour experts are – as opposed to direct solo choice architects – mostly co-, contra-, and even more indirect choice architects. In light of those roles and the networking, transferring, and translating activities they imply, behaviour experts are therefore better understood as knowledge brokers.

Adopting a knowledge brokerage perspective can help to understand the typical challenges behaviour experts face. As new kids on the block, a crucial challenge for them lies in becoming part of the established order. They have yet to prove their added value in a rationally bounded policy system that is not too susceptible to new influences. However, their precariousness is not only a result of their novelty but also of their identity as knowledge brokers, always operating 'at' and 'in between' organizational boundaries. Knowledge brokerage tends to be a relatively 'boundaryless' and 'invisible' professionalism, characterized by role conflict, role ambiguity, a lack of organizational recognition, and a lack of career pathways (Chew et al. 2013). These issues are likely to affect behaviour experts as well. In this light, it is constructive to point out some of the strategies that knowledge brokers employ to cope with their precarious situation, and even using it to their own advantage as a means towards greater flexibility and autonomy (Chew et al. 2013). Thus far, several of such strategies have been identified, including: relying on collective forums and peer support networks; relying on additional boundary-spanning actors; ensuring dual participation from both science and policy; ensuring dual accountability; creating 'boundary objects' that connect science and policy; facilitating co-production; and strengthening internal meta-governance and capacity building (Chew et al. 2013; Hoppe 2010; Rouw 2011). Learning about these kinds of strategies and translating them to the context of the behavioural policy practice may help behaviour experts in overcoming knowledge brokerage-related challenges.

**Table 3.4: Key practices of behaviour experts**

| *Key practice* | *Choice architecture* | | | *Analysis* | *Capacity building* |
|---|---|---|---|---|---|
| | *Solo choice architecture* | *Co-choice architecture* | *Contra-choice architecture* | | |
| *Activities* | Making and evaluating interventions | Setting up and managing the making and evaluating of interventions | Regulating commercial choice architectures | Analysing policies and policy targets | Educating and persuading fellow bureaucrats, and producing knowledge |
| *Goal* | Effectuating | Organizing | Regulating | Understanding | Awareness-raising |
| *Role* | 'Choice Architect' | 'Network Node' | 'Inspector' | 'Analyst' | 'Ambassador' |
| *Scale level* | Micro | Micro | Micro | Meso | Macro |
| *Primary target* | Citizens and businesses | Internal and external policy producers | Businesses | Citizens and businesses | Internal policy producers |
| *Example* | Redesigning school canteens to stimulate healthy eating | Bringing together a network of partners to run a field trial on anti-loitering | Pressuring commercial travel businesses to remove harmful pre-ticked boxes | Observing citizens in their kitchens to understand their waste sorting behaviour | Teaching military units to use behavioural insights in the field |

To conclude, the knowledge brokerage perspective opens up new avenues for future research. It may be interesting to look further into the role of context when brokering behavioural science. As the chapter has shown, behaviour experts operate in a wide range of policy environments which differ greatly in terms of which actors are involved, what tasks are being executed, how responsibilities are divided, what kind of policy problems are being tackled, and how much familiarity there already is with behavioural policy approaches. These differences are likely to affect the role that is asked of behaviour experts. To give just one example, in relatively coherent fields of governance it may be possible for behaviour experts to concentrate on helping a small

set of key policy actors devise practical applications. The focus then lies on transforming abstract ideas into concrete outputs. However, in the highly disjointed field of governance in which policymakers are merely meta-governing, it may be more necessary for behaviour experts to work on building stronger science–policy linkages and smoothing the flow of information to and between all relevant actors. The focus then lies more on the transferring aspects of knowledge brokerage. Research on how these kinds of contextual differences call for distinct ways of brokering behavioural science seems worthwhile. Furthermore, it may be interesting to shed light on the dynamics between knowledge brokering and institutional innovation more generally. We have seen that the attempt to set in motion a deep institutional change, in this case the building of the Dutch behavioural state, has materialized into a specialized group of public professionals whose practices tie in perfectly with typical knowledge brokerage work. This leads to further questions about how important knowledge brokering is for institution-building and whether other newly emerging professions in governments are 'knowledge brokers in disguise' too. Despite knowledge brokering being a naturally less visible, backstage, or perhaps more accurately, 'between-the-stage' phenomenon, it may well be a core driver of institutional change.

*Chapter 4*

# Localizing the Local Behavioural State

*Policymakers in Dutch government are increasingly making use of behavioural insights in the policy process. These insights have initially been welcomed at the central state level. However, they seem particularly promising at the local level. Behavioural policy seeks to change behaviours by redesigning the immediate environments surrounding citizens, and it is the local governments that have control over those environments. In light of this promise, this chapter explores the recent emergence and institutionalization of behavioural expertise in local government. The practices of local behaviour experts are analysed through three dimensions: position, practices and politics. The findings demonstrate that local behaviour experts, while still in an experimental stage of development, are well aware of and critically engaged in various facets of behavioural challenges. Through their practices, they strategically cope with scarce resources, resistance and competing institutional logics.*

---

Chapter 4 is based on an original research article by Joram Feitsma in *Bestuurswetenschappen*. This article has been translated from Dutch to English language.

Feitsma, J.N.P. (2019). Meters maken: gedragsexperts in het lokaal bestuur. *Bestuurswetenschappen, 73*(1): 5-29. doi: 10.5553/Bw/016571942019073001002

## 4.1 Introduction

The application of behavioural science in policy is high on the strategic agenda of Dutch government (e.g. BIN NL 2017). Under the intellectual guidance of The Netherlands Scientific Council for Government Policy (WRR) (Tiemeijer et al. 2009; Jonkers & Tiemeijer 2014; WRR 2017), a collective investigation has begun exploring how behavioural insights can help to make smarter and more effective policy that is sensitive to actual citizen behaviour. An integral part of this effort involves probing and questioning the level at which behavioural expertise might be most fruitful. Until now, the overwhelming bulk of attention has been given to the central government context (BIN NL 2017; Schillemans & De Vries 2016b; Feitsma 2019; Feitsma & Schillemans 2019), where a regard for strategic policy innovations exists naturally. In this chapter I will shift the focus to a lesser explored, but possibly (more) promising, locus: *local* government. Behavioural insights are essentially concerned with acknowledging and exploiting the ways in which citizens are influenced by their immediate environments. Local officials in particular have both knowledge of and a solid grasp of these environments. Although there are few studies into local behavioural public policy, various developments are causing interest in the field to rise. A recent example (see Figure 4.1; @Marielle1972 2018) comes from the City of Rotterdam, which collaborated with behavioural scientists in an effort to prevent cyclists riding on the wrong side of the Erasmus Bridge. Communication signs were inserted on the cycling lane to encourage the 'booing' of these so-called 'ghost cyclists', thereby tapping into the human capacity to attune to and process (negative) judgements from the social environment. This *shaming* technique resulted in a significant reduction of ghost cyclists per day (Kooyman 21 August 2018). This chapter explores the emerging institutional practice behind these types of innovative behaviour change interventions. It describes the early development of behavioural expertise in Dutch local government. The research question is: *what can we learn from practices of pioneering local behaviour experts about the opportunities for developing local behavioural public policy?*

## 4.2 The rise of the steering state

The Dutch developments regarding the use of behavioural science fit within an international policy agenda, which had its genesis in Anglo-Saxon countries (Leggett 2014; John 2018; Whitehead et al. 2017). The global 'behavioural turn' is due, in part, to the publication of *Nudge* (2008) from the American scientists Richard Thaler and Cass Sunstein. This book has made insights from the new school of behavioural economics accessible to a wider audience. This *new school of behavioural economics* emerged in the 1970s and 80s out of the desire of economists to make their models

better attuned to the ways in which human behaviour deviates from the principle of rationality (Leggett 2004). Fuelled by insights from cognitive psychology (e.g. Tversky & Kahneman 1974), behavioural economists identified the predictable patterns in which people deviate from the *homo economicus* ideal (i.e. the purely rational decision-maker). In addition, a rich catalogue of cognitive biases has emerged, including the *halo effect*, *sunk cost fallacy*, *optimism bias, confirmation bias* and *authority bias* (Kahneman 2011; see Benson 2016 for a parsimonious overview). Together these biases demonstrate how people are – driven by their fast, automatic, *System I* thinking (Kahneman 2011) – more susceptible to influences from their physical and symbolic environments than commonly assumed.

**Figure 4.1: Shaming on the Erasmus Bridge**


The power of *Nudge* lies in its ability to translate the new behavioural economics insights to a broader policy audience. In doing so, these insights can be explored and applied to policy problems. Here we enter the domain of the so-called 'nudges', which

shape the 'choice architecture' (Thaler & Sunstein 2008) in such a way that certain practices can elicit desired behaviour without regulatory or financial coercion. Examples of such subtle nudges are omnipresent. In the Netherlands, for instance, take the default change for organ donor registration (from opt-in to opt-out), the attractive placement of healthy products in school canteens, and the hiding of tobacco products in supermarkets. The Behavioural Insights Team (BIT UK) in England, launched in 2010 as a strategic unit under the Cabinet Office, serves as a trendsetter for these types of nudge interventions, and also for the way in which they conduct and disseminate their resulting evaluations (Haynes et al. 2012; Service et al. 2014). BIT UK's empirical-scientific approach seamlessly connects to a broader trend of evidence-based policymaking (Parsons, 2002; Rouw, 2011).

The emergence of behavioural public policy has been the subject of a polarized and puzzling academic debate. On one side we see staunch and enthusiastic advocates, who above all view behavioural policy as a promising if not necessary avenue towards more effective and less coercive policy practices (e.g. Thaler & Sunstein 2008; John 2018). On the other side are critics, who have voiced doubts about the epistemological and moral agenda of behavioural policymaking. Major moral critiques hold that behavioural policy is manipulative and undemocratic as it essentially tries to steer citizens from a distance by tapping into aspects of their unconscious decision-making (e.g. White 2013). An epistemological critique holds that behavioural policy is underpinned by a narrow focus on cognitive psychology and experimental methodology. It is also associated with a naïve social engineering mentality (e.g. Lodge & Wegrich 2016); a new form of 'greedy governance' (Trommel 2009) which foolhardily seeks to bend society to its will. In addition to such advocates and critics, there is also a group of sceptics who are concerned with how valuable, sustainable and innovative this 'megatrend' really is (e.g. Bonell et al. 2011b). Sceptics view the apparent insights stemming from *Nudge* as an inflated institutional fantasy; an ephemeral hype with grand ideas but ultimately little substance.

This confusing mix of conflicting analyses – i.e. the behavioural state as utopia, dystopia and illusion – makes the societal implications of this state uncertain. In an attempt to further unravel this debate, this chapter is devoted to a more preliminary, descriptive question: how has this trend actually arisen and what concrete institutions has it produced? In other words, this chapter studies the institutionalization of the behavioural turn. It does so by taking the perspective of the early adopters promoting this turn within the context of Dutch government. By studying local behaviour experts and the emerging institutional configurations linked to them, I explore how behavioural policymaking occurs in practice and interrogate the associated opportunities and risks of such an approach. This study is less concerned with the

end products of the behavioural state, such as the experiments, nudges and behavioural interventions themselves. Instead, the focus of this study is on the local public officials who profile themselves as 'behaviour experts' and their progress (their institutionalization trajectory) as they search for a stable and legitimate place in what is already a crowded institutional space.

Over the past decade there has been significant interest in the application of behavioural insights in the Netherlands. This is reflected both in a series of influential advisory reports (e.g. Tiemeijer et al. 2009; Jonkers & Tiemeijer 2014; WRR 2017) as well as the establishment of BITs and behavioural networks across a range of government agencies and departments. Units have for instance been launched at the Dutch Tax and Customs Administration, the Authority for Consumers & Markets (ACM) and the Ministries of Economic Affairs and Climate Policy (EZK), and Infrastructure and Water Management (IenW). In addition, an interdepartmental Behavioural Insights Network Netherlands (BIN NL) has emerged, adopting an agenda-setting and coordinating role (BIN NL 2017). Dutch public administration and policy scholars have also begun exploring the implications of the emerging steering state in a special issue titled, 'Behavioural insights in policy and public administration' in *Bestuurskunde* (Schillemans & De Vries 2016a). This special issue investigates various examples of behavioural policy, pinpoints possible ethical and moral issues, and delineates opportunities for further institutionalization of the field. The bulk of scholarly attention, however, is devoted to specific behavioural interventions, and not to the rise and professionalization of behaviour experts. The majority of these studies are largely concerned with central state level practices. This orientation may be a consequence of the top-down fashion in which behavioural insights emerged, initially via the WRR report *The human decision-maker* (Tiemeijer et al. 2009) and the resulting interdepartmental knowledge platform.

### 4.3 Local behavioural expertise as *terra incognita*

A danger of the abovementioned scholarly orientation is that it gives the impression that Dutch behavioural expertise is found mainly in the central government and that such practices are conducive with centrally orchestrated, top-down policymaking approaches. Such an impression would overlook numerous novel and interesting developments regarding the application of behavioural science that take place at the *local* level (for instance in municipalities, schools, corporations, healthcare facilities and umbrella organizations). A range of behavioural initiatives and practices have emerged at the local level, including the 'Nudge Network Utrecht' (City of Utrecht), 'Behaviour changer' (City of Hoorn), 'Nudge expert' (City of Deventer), 'Adviser behaviour change' (City of Enschede), 'Nudge consultants' (GGD Friesland) and a

'Behavioural Insights Group Rotterdam' (City of Rotterdam). These initiatives are not merely the by-products of the central state's knowledge agenda. Instead, these practices find their roots mostly in concrete localized policy challenges that demand innovative solutions.

In fact, behavioural expertise may be particularly fruitful at the local policy level. This is because it involves the psychologically-informed tweaking of the physical and informational environment (i.e. choice architecture) surrounding citizens (Whitehead et al. 2018). The position that local behaviour experts have, relative to their central state counterparts, is quite advantageous. Particularly with regard to their proximity and understanding of what happens 'on the ground.' They are positioned 'closer to the citizen' and have a better view of and control over citizens' immediate environments and their resulting behaviours. While direct contact between the central state and citizens is relatively limited and relegated to the informational sphere, with letters and campaigns as primary means of contact, the opportunities for direct engagement between the municipal government and citizens are more expansive. This is because municipalities are responsible for the design of local infrastructure and physical public spaces. All sorts of behavioural norms apply to these spaces, for instance related to speed limits, parking fines, or requirements for local waste management. The government's task of ensuring compliance with these norms lends itself well to choice architecture at the local level.

Against the background of developments over the past decades in which the dominant policy style has moved *from government to governance,* and *from welfare state to participation town,* behavioural expertise also seems promising at the local level. Increasingly, central state power has shifted toward local governments as *the first government* (Paardekooper et al. 2013), and increasingly citizens themselves are held accountable for social issues (Ossewaarde 2007). The recent decisions to decentralize essential tasks (healthcare, employment, and youth care) in the social domain are illustrative of this power shift. As a result, local public administrations take on an increasingly important role as choice architects. It is, however, not completely clear whether this role is assumed consciously and/or competently, and whether the necessary resources, knowledge and skills exist to facilitate such a role. Innovative behavioural expertise – and its associated promise of smart, rigorous and effective policy – would seem to be of value as local administrations navigate their new and expanded responsibilities.

Local behavioural expertise thus seems like the *terra incognita* of behavioural public policy: although there are plenty of local opportunities to put behavioural insights to use, local behaviour experts are a relatively small and hidden community which is

often missing from public, political and academic debates. To stimulate such debates, and to be able to delineate both the opportunities and limits of local behavioural policy, this chapter will explore the work practices and institutional arrangements of local behaviour experts. Its scientific relevance lies in uncovering the little explored terrain of local behavioural expertise – and accordingly, in thinking through where the real opportunities of this trend lie. In a broader sense this chapter also offers insights into new developments concerning the interplay between science and policy. Thus, it sheds new light on the classic debate about the role of scientific knowledge in policy practice (Lindblom 1959; Lindblom & Cohen 1979; Parsons 2002).

My exploration of local behavioural expertise will be underpinned by a model with three dimensions – position, practices, and politics – that can be used to characterize local governance. The dimension *position* refers to how local governmental actors are organized and positioned. The dimension *practices* concerns the daily activities of local policy actors, the theories, methods and instruments they employ, as well as the results they achieve. The dimension *politics* relates to the ideological beliefs, values and ethics that underpin local policymaking. It concerns the degree to which local governance leads to 'fuss', conflict and struggle, and how controversial issues shape local policy processes. These three dimensions are derived from Boogers and Schaap (2007), who reflect on the way in which access to internal and external resources and partners (*position*) shape local policy, as well as from Barber (2013), who points out how the type of task (*practices*) and political-administrative culture (*politics*) determine local policy. The model is also closely related to the work of Scherpenisse et al. (2016), who breakdown the strategic issues involved in the establishment of behavioural practices into choices about positioning, projects, performances and professionalism (*position* and *practices*). These three dimensions – which are taken as sensitizing concepts – are detailed in Table 4.1 (derived from Boogers & Schaap 2007; Barber 2013; Scherpenisse et al. 2016).

**Table 4.1: Three dimensions of local behavioural practices**

| *Dimension* | *Position* | *Practices* | *Politics* |
|---|---|---|---|
| *Aspects* | - Internal positioning<br>- Available capacities<br>- External knowledge infrastructure | - Tasks and application areas<br>- Activities, outputs and outcomes<br>- Theories and methods | - Political-administrative culture<br>- Struggle, conflict<br>- Resistance |

### 4.4 Methods

In order to study the rise of local behavioural expertise in detail, and at the same take account for its varied manifestations, I have done a 'step-in step-out' (Madden 2017) variant of ethnographic fieldwork (Rhodes et al. 2007). Over a period of 17 months (November 2016 until March 2018) I immersed myself in the daily environments of various local behaviour experts. I studied them in their professional habitat for a relatively long time, not only querying them from the outside ('frontstage') but also observing them from within ('backstage'). This ethnographic approach helps to capture what 'really' happens within organizations (Van Hulst 2008). More specifically, I have used three methods:

- Participant-observation on 11 different occasions, totalling 19 hours;
    - 4 observations in which I attended an internal educational meeting;
    - 4 observations in which I attended an intermunicipal knowledge exchange;
    - 1 observation in which I helped carrying out a field experiment;
    - 1 observation in which I contributed to a behaviour change course for a municipal department;
    - 1 observation in which I helped a BIT in thinking about the internal organization of behavioural expertise.
- Semi-structured interviews with 15 local actors, predominantly behaviour experts:
    - 11 officials at 8 different municipalities (Utrecht, Amsterdam, Rotterdam, The Hague, Enschede, Deventer, Leeuwarden and Hoorn);
    - 2 officials at intermunicipal umbrella organizations (The Association of Netherlands Municipalities (VNG) and Divosa);
    - 1 alderman;
    - 1 school director.
- Document study of memos, email correspondence and online content.

Initial remarks, insights and analyses on the basis of the observations and document study have been processed into fieldnotes. These fieldnotes have been carefully and closely scrutinized in order to be presented in this chapter. The semi-structured interviews have been recorded, selectively transcribed, and, together with the fieldnotes, coded in an iterative process using the sensitizing concepts from Table 4.1. These concepts offered a generic structure to analyse the empirical data and make sense of typical perspectives and salient similarities/disparities in the field. Within this structure – with *position, practices* and *politics* as main codes – various sub-codes

have been distilled from the empirical data. For instance, when analysing the empirical material in relation to the main code *position,* certain topics would repeatedly arise – like 'central unit versus network approach' – which therefore have been taken as sub-codes. The analysis presented here has been composed using these main and sub-codes.

Within this context, a question of utmost importance is: What actually makes someone a 'behaviour expert'? This question is not unequivocally answered given the hybrid history of behavioural public policy. To begin with, in an implicit sense, most social scientists as well as many public officials with knowledge of the local citizenry might be regarded as 'behaviour experts'. The use of behavioural insights on the local level has a long tradition. This includes, for example, use of social psychological intervention techniques in domains such as municipal informing, youth care, enforcement, waste management and local health care (GGD). However, over the past decade, the Nudge trend has certainly given a new impulse to the use of behavioural science: an impulse that is largely based on a distinctive source of behavioural insights, i.e. the new school of behavioural economics and its focus on biases and heuristics. It is this recent trend that this chapter focuses on. I look for those who identify themselves as its leaders: the self-professed 'Nudge experts' and 'Behavioural Insights Teams'. I define 'behaviour experts' as local actors who explicitly and structurally make use of new behavioural economics insights, and present themselves as champions of this new trend to the outside world. As a new professional community, these champions form the most direct representation of *Nudge*, and therefore a relevant study object to better grasp this trend.

Local behaviour experts constitute a fairly new and fragmented group. Identifying and selecting respondents and opportunities for participant observation presented a considerable challenge. A lack of available public documentation led to a pragmatic sampling strategy. I started by drawing from my own existing network, getting in contact with a behaviour expert who was involved in the 'Nudge Network Utrecht' as well as in an intermunicipal G6 gathering concerning behavioural insights. This initial contact led to several interviews, participant-observations and invitations for other meetings, which accordingly – using the snowball technique and doing desk research – helped me in mapping out and identifying additional local behaviour experts. Wherever in the field I have found structural activities in using behavioural expertise, I have made visits and talked to experts – except for those in Joure, Eindhoven and Tilburg which came into view only later during the research process. The research sample mostly includes behaviour experts who are active in relatively large municipalities in metropolitan areas. This ties in with the idea that resource-rich local governments are more prone to experimenting with policy innovations (Shipan

& Volden 2012). Besides studying behaviour experts active in different regions, I have also studied a varied set of local governance actors working both within municipalities and also in other institutions.

### 4.5 Explorations in the local Behavioural Insights landscape

The local behavioural landscape is slowly maturing and evolving. In the past decade behavioural units, networks and functions have been launched in diverse (inter)municipal and local social organizations. The municipalities in Western Holland seem to have taken the lead in terms of these developments. Although local behaviour experts present a united front in their strong rationalist message – i.e., aspiring to make more effective policies by using novel behavioural insights – they are a relatively fragmented group. They are active in a broad range of policy areas, for instance public health, communication, household waste management, regulation and enforcement, employment, and debt and poverty. Heterogeneity also exists in terms of professional backgrounds, varying from respondents who are relatively new to public behaviour change to those who have significant first-hand experience and/or academic knowledge on the matter. Respondents with more advanced expertise have for instance followed behaviour change masterclasses or have undertaken training or academic studies in a directly relevant field, such as social psychology. The City of Rotterdam's behavioural unit offers one example of a unit with relatively advanced expertise. One of the leaders of the unit has a background in health psychology, while the team members specializations range from health sciences, psychology, and anthropology (BIG'R Website 2018).

Given the field's emergent and heterogeneous character, behaviour experts are still trying to map their own field and gain an understanding of who is actively involved in using behavioural insights. A lot of connections emerge in a coincidental and incremental manner, and people are working in 'silos'. One respondent notes: 'We increasingly come to know each other through coincidental liaisons. Through word of mouth you'll hear: "That one's also working on it, that one has done a project." Those kind of things. And still nearly every week I am talking to new colleagues who are also involved. Yesterday even.' There are attempts to create more structural connections between and within municipalities. An intermunicipal G6 gathering has for instance been launched, in which experts in the Cities of Utrecht, Amsterdam, Eindhoven, Tilburg, Rotterdam and The Hague are exchanging knowledge and experiences. I have attended several of these G6 gatherings. These participant-observations proved to be extremely meaningful: in the presentations that behaviour experts would give to each other and through the resulting conversations, a deeper understanding of the broad range of aspects of their work came to light. Observing

this intermunicipal gathering also helped in comparing similarities and differences between the varied individual approaches.

To further analyse the local behavioural landscape in more detail I will zoom in on the three dimensions of position, practices and politics. This analysis is inextricably linked to a reconstructed observation of one of the aforementioned intermunicipal G6 meetings, i.e. the third G6 gathering in March 2018 in the town hall of Rotterdam. This meeting took place there for a reason, as at the moment of research Rotterdam is currently featuring the by far largest local behavioural team. I present my observations of this G6 gathering in the form of scrutinized fragments which stem from initial fieldnotes. These fragments have been edited and adjusted for readability and scanned for relevance in light of the sensitizing topics (see Table 4.1). The focus on these specific occasions helps to capture the local behavioural practice with its associated particularities from up close. At the same time, where relevant, the analysis is also supported with other empirical data in order to account for how local behavioural expertise manifests in the field.

***Position***

The analysis first looks into the *positioning* of behavioural expertise at local level.

> The chairman opens the meeting. The experts from Amsterdam and Eindhoven are unable to attend. On the agenda are four presentations, preceded by a short introduction by a member of 'Behavioural Insights Group Rotterdam' (BIG'R).
>
> The BIG'R member talks about the type of projects they are occupied with, like parking, public health and nightlife-related nuisances. 'As a municipality it's the more practical matters that we get to deal with. And from officials we often get the question: "Can't you do something with that?"' The BIG'R member notes that there are different approaches being employed by municipal behaviour experts. Some go for hiring external bureaus. Others tackle things very thoroughly by themselves. He makes a call to share with one another what works, 'So now here's the question: why can't we share some things with each other so that we don't have to do this from a solitary position? And if something doesn't work? Let it then fail spectacularly so that others don't have to make the same mistake.'
>
> The BIG'R member notes how his team has received a budget of 1 million for four years in order to set up a collaboration with Erasmus University Rotterdam. 'We buy two researchers from Erasmus University. And these

do not stand at a distance', he says as he points to the two people next to him. 'They are right here. More than a year ago we started with 4 A4 sheets, searching for projects, and now we're with 25 part-time and full-time people. We've got an infrastructure, a number of topics, and are ready to make head-way.' The BIG'R approach has since been further formalized, each time working towards an end product which can vary from ad hoc collegial consultation to larger projects consisting of extensive behavioural research.

It is the behaviour expert from The Hague's turn to present. She reports that The Hague is still at a start-up phase when it comes to the municipality wide use of behavioural insights. The approach has thus far been to work with externally hired bureaus. 'But we increasingly want to move to a situation where we have access to that knowledge in-house.' It has been a challenge to bring behavioural expertise 'to a higher level. Thus far there's been too little knowledge, too little time, too little means and too little money to get started with this.' So far she has mainly invested in building a network, starting a collaboration with the local university, and collecting cases. She is also thinking about how behavioural insights might be applied across the municipality in the future, possibly with the help of an 'expert unit'. Currently, a group of circa twenty people, employed at different policy levels (managers, policymakers, implementers) is interested to work on this topic.

Utrecht's behaviour expert starts by mentioning that Utrecht doesn't have a formal BIT. He and a colleague have however organized all sorts of associated activities in their own time. In ten pairs, a group of local public administrators are now for instance following behavioural change masterclasses. Also, they would like to create one full-time position for this theme to 'keep track of what's happening'. The focus thus lies on a decentralized approach. 'I want to outsource this and ensure that departments themselves get started with this. Preferably, I'd make myself redundant and colleagues would take action themselves. Within my own department this has already happened quite a bit. That's precisely what I want actually. I don't know anything about the [policy] content anyway. [It's] a nice, natural way of working.

One of the experts from Rotterdam asks his colleague from Utrecht: 'Have you already got a lot of people with expertise? And how do you ensure that people really have sufficient intellectual capacity to get started with projects? Our experience is that it's easy to trigger people, but to really

> have them make the translation to policy that leads to behaviour change… that requires a bit more than one short course.' The expert from Utrecht responds: 'That depends on what you mean with "expertise". If you mean people with some degree of expertise, those are in ready supply.. And with regards to our courses I've really got the idea that we have a number of colleagues who remain enthusiastic about this and that there's a degree of anchoring. But true "experts"… we don't have any of those.'

The fieldnote documents numerous interesting aspects. First, it shows that there is an explicit and significant appetite for activities associated with positioning, network building and further institutionalization. Steps are being taken to further align the different units, networks and initiatives and to anchor them in the existing order. The recurring meetings as part of an intermunicipal G6 gathering is a striking example of this. But also at the intra-organizational level, behaviour experts are experimenting with suitable formats for the application of behavioural insights. Diverging approaches become visible. At present, it seems that most local behaviour experts are following a 'network approach', in which existing behavioural practices operate autonomously yet are also part of an interconnected network. In this context, municipal behaviour experts are mostly active as (temporary) 'encouragers'.

The dominance and inclination toward a network approach partly seems to be a consequence of the explorative stage in which most local behaviour experts find themselves operating within. The focus of this stage lies in surveying the field and attempting to gain an overview of current activities regarding the intra-organizational use of behavioural insights. This tendency is exemplified in the activities of the behaviour expert from the Hague, who notes that she brought together a 'team with interested persons' which recently gathered for the first time. At present, the apparent structure remains relatively informal. Next to this network approach, there are also modest attempts to formalize and concentrate behavioural expertise. These attempts use the BIT UK as a model as it was initially established as a formal and autonomous unit. The City of Amsterdam for instance employs a municipality wide 'Behaviour Expert' and the City of Enschede hosts an 'Adviser Behaviour Change' who uses behavioural science on a full-time basis. In terms of financial resources and capacities, the relatively ambitious approach of the City of Rotterdam stands out, which in collaboration with Erasmus University Rotterdam has launched a formal unit with 25 members.

Most local behaviour experts note that the scarcity of available resources poses a particular challenge. Deep behavioural expertise is often lacking, and time and resource constraints make it difficult to conduct high quality projects. Often

behaviour experts are expected to do their behavioural work in addition to their regular tasks. An expert told me during a field observation that it can be demoralizing when others have to be motivated constantly and when no further recognition or formalization occurs. However, the more formalized and centralized units are still relatively minor characters with little capacity if considered in terms of the broad swathe of public service activities. In this light, the advantage of a decentralized network approach is that local experts are able to latch on to existing projects, as well as their associated capacities and budgets. One behaviour expert mentioned that when he would connect to the substantive projects and interests of budget or project managers (in his case for instance relating to decreasing traffic jams) there would be 'an abundance of resources'.

Local behaviour experts collaborate frequently with external partners. They bring in knowledge and competencies through their interactions with scientists, consultancy bureaus and international policy agencies. Important partners that are conducting research and teaching courses are, among others, Duwtje, Dijksterhuis & Van Baaren, BIT UK, Erasmus University Rotterdam, and Utrecht University (I have also been involved in local collaborations as a Utrecht-based researcher). As the fieldnote suggests, there is a distinct desire to establish and develop more in-house expertise, for instance by having municipal officials attend masterclasses. At the same time, legitimate questions arise as to whether these initiatives are actually worthwhile, whether such efforts are capable of capturing the requisite knowledge, and whether more intensive investments should be made.

In short, the application of behavioural expertise occurs in a somewhat diffused manner, scattered over loosely coupled projects. In addition, the application is partially externalized, given the intensive collaboration with external bureaus and universities. Where behavioural expertise is more concentrated and formalized, an increased potential emerges. A municipal BIT – like BIG'R in Rotterdam – can function as a central arena where knowledge and resources convene. It gives a tangible and concrete institutional shape to the abstract idea of 'the use of behavioural insights'. Furthermore, it contributes to the internalization and anchoring of behavioural expertise. And yet, some behaviour experts prefer the more diffuse and less visible network approach. According to such experts, this approach is more suitable as it capitalizes on pre-existing initiatives and, as one respondent noted, 'a stronger sense of ownership emerges. If there's a central unit, then "it's their thing" and I don't have to do anything with it anymore, so to speak.' This issue was concisely worded as the dilemma or trade-off between 'traceability versus ownership'.

The way in which behavioural expertise gets institutionalized in the different municipalities, and its resulting potential, is inextricably linked to the unique aspects of the local environment. For example, proximity to a university with expertise and ambitions in the study of behaviour change seems to open up a range of different opportunities. This is observed in the case of BIG'R closely collaborating with Erasmus University Rotterdam, or the co-produced Living Lab in Utrecht. Municipalities in larger urban areas are better placed and have greater capacity to capitalize on and institutionalize behavioural expertise, as they usually have a greater appetite for strategic innovations, possess more in-house expertise, and have better connections to external knowledge partners. Also the influence of the local political-administrative culture becomes apparent in how behavioural expertise is organized. The stereotypical Rotterdam 'can do' mentality is seemingly reflected in the launch of BIG'R, a unit which was able to build capacity and institutionalize relatively quickly.

***Practices***

The analysis now shifts to a focus on *practices,* for example the working areas, approaches, and day-to-day activities of local behaviour experts.

> The expert from The Hague talks about the behaviour change cases that exist in her municipality. Those appear to be in abundance: behavioural insights have been used in various policies relating to, among others, underground collection units, retirement, burglary, mobility mentoring, and healthy bread bins. There are also scheduled projects regarding bicycle parking, outgoing letters, illegal house rental, no-shows, handling stolen goods, pickpockets, and domestic animal waste in public spaces.
>
> The Rotterdam-based expert, also a BIG'R member, recounts an (in his eyes) 'atypical' case. Concerning a newly opened swimming pool in the city it was expected that its publicly accessible entrance of the pool might become a site of anti-social behaviours (like loitering, licentious behaviour, graffiti, drug use, homeless persons sleeping) without surveillance being possible. Such anti-social behaviours could scare off pool visitors or subject them to unpleasant experience. The risk of anti-social behaviour has to do with the pool's location on the first floor, and that the entrance doesn't give immediate access to it. BIG'R has looked how the risk of anti-social behaviour could be reduced. BIG'R project members put an internal scientific employee on this case, visited the location a few times for observation, and have entered into conversations – for instance with the architect, operator and the area manager. Accordingly, they brainstormed about possible interventions stemming from the theory of 'defensible

> space', which emphasizes aspects of natural surveillance, visibility and continual flow in regulating social behaviours in public space (BIG'R Website 2018). From this activity, BIG'R advised to invest in making the entrance space more transparent and creating a sense of 'human presence'. Important suggestions – that have now been realized – were to work as much as possible with glass material in terms of architectural design. The lift shaft as well as parts of the entrance were redesigned and made entirely out of glass – an impactful but also an 'expensive mutation'. BIG'R is also further considering how behaviour can be triggered by working with visuals, light, sounds, and smell. The Rotterdam-based expert explains: 'If we look at the senses we've got… seeing, hearing, smelling – tasting might perhaps be less interesting for now – then it seems that there are all sort of sensory routes through which behaviour can be steered.' Ideas for instance are to play 'swimming pool sounds', to portray images of humans on the wall, and to work with lighting that mimics the reflection of water – so as to create an inviting 'atmosphere' that is reminiscent of the flow of the pool (BIG'R Website 2018). Another idea was to install a fragrance system with the smell of chlorine. These ideas have been shelved for now, as no immediate reason to use them has become apparent yet: since the opening of the pool, the dreaded nuisance has not emerged.

What first emerges in this fieldnote is the richness of identified cases for behaviour change. Some tend to be 'very banal things', like tackling the relatively minor public disturbance that is the result of incorrectly parked bikes. According to behaviour experts, municipalities are still unable to steer citizens in the desired directions in many of these types of matters. Experts in the area of communication, waste collection and public health for instance mention that a predominantly rational persuasion approach produces insufficient behaviour change. Further, experts in the area of regulation and enforcement, and employment see the predominantly coercive approach as largely ineffectual. New behavioural economics insights are embraced as a promising and innovative perspective to move beyond these shortfalls. Although this embrace extends to the breadth of policy areas, certain popular themes exist that include the stimulation of exercise and healthy eating, reducing no-shows at the municipal counter and reducing public littering.

While local behavioural practices possess a *wide* area of application with various opportunities for choice architecture, a question remains as to whether such opportunities can be *thoroughly* seized. At first glance, such worries seem superfluous; if anything, local behaviour experts emphasize the importance of working thoroughly

and rigorously. This is reflected in their modus operandi, that is, their intention to tackle and analyse issues scientifically, systematically and comprehensively – on the basis of powerful causal evidence of 'what works' (cf. Lindblom's (1959) *rational-comprehensive model*). This rationalist, evidence-based narrative can be recognized in actual practice, particularly in the diversity of field experiments that have been run. An Amsterdam-based expert, for instance, ran a range of experiments to test the effects of social norms in traffic communication which was intended to discourage red light negation. Utrecht University and the 'Nudge Network Utrecht' together ran experiments on whether redesigns of the canteen environment could stimulate water drinking and standing (versus sitting). The City of Enschede, together with a consultancy, ran a letter experiment to increase the compliance of debt holders. In a letter to non-paying debt holders various nudge techniques were applied, including presenting a social norm, simplifying and visualizing information, and reframing information to avoid negative emotions (e.g. refraining from using the term 'backlog'). This resulted in a relative increase of 30% more debt holders who took action regarding their financial situation (Janssen et al. 2017).

However, following the evidence-based approach dogmatically or steadfastly appears a difficult task. An interviewee states: 'I'd like to apply the knowledge [behavioural insights] more often and more thoroughly than what I'm doing now.' In practice, local behaviour experts have the experience of dealing with a tension between scientific rigour on one hand and political and practical feasibility on the other. They possess, as the previous section has articulated, a limited amount of time, resources and expertise. They also deal with dynamic choice environments in which the outcomes of policy actions are not easily measured or made visible. Such limitations in practice often require compromises and trade-offs between rigour and feasibility.

This idea of compromise is made explicit in the observation that local behaviour experts partially replace their rigorous practice based on experimental methods (cf. Aristotle's' notion of *episteme;* see Parsons 2002), with a 'softer' approach grounded in a mix of other methods like expert consultations, interviews, surveys, observations, professional reasoning, creative thinking and practical intuition (cf. Aristotle's notion of *phronesis;* see Parsons 2002). This softer approach has led to the design and implementation of numerous innovative and original behaviour change interventions. An Utrecht-based behaviour expert mentioned that he had been involved in the design of a range of interventions. Such interventions included the design of attractive and colourful plates intended to promote the consumption of smaller portions of birthday sweets in primary schools; 'growth cards', a series of cards that youth health nurses can give to parents during their consultation, with a

brief and visual explanation of how children can be involved in daily activities to stimulate their cognitive development; and design concepts that playfully encourage healthy behaviour within the new Utrecht railway station area, for instance a 'playful route' with a range of gaming elements in public space.

A certain degree of adherence to scientific rigour is thus exchanged for a faster, softer and more pluralist approach. Here, the logic of scientific rigour meets a logic of rapid judgements based on practical experience and common sense (Lindblom & Cohen 1979). An Utrecht-based expert articulated a preference for a 'pragmatic' approach, based on an ongoing learning process through quick 'pilots'. Ultimately, he preferred to work 'practice-based' rather than 'evidence-based'. This preference was not contentious, rather it was a matter of 'simple everyday practice' in which conducting scientific study was not always an option or necessary. A similar approach is also visible in the swimming pool case from the field note. BIG'R didn't need to run laborious and time-consuming experiments, instead it 'confined' its process to the study of salient literature, observations and conversations. Using this preliminary work as its foundation, the team then brainstormed about smart behaviour change interventions. A stronger example of compromise regarding the reliance on scientific rigour emerges from an interview with a school director. The school director explicitly distanced himself from a hard scientific logic. In the context of an infrastructure project at the school, the director opted to use behavioural insights from a more pragmatic logic. Acting on the basis of 'what works' was largely a matter of 'intuition' (also see Feitsma 2018a):

> How great would it be if we'd known that the colour 'red' triggers people's motivation to take the stairs, for instance? How great would that be? But in a different way, that's also a utopia and you'd have to experimentally assess that. You'd really have to do it: the experiment. And on a small-scale that may be possible but on the large scale it's mostly still about intuition though. Maybe yeah it's the case that… (…) you'll do some literature study (…). But even then it's a cost-benefit story. And then I think: "Well, there are a lot more things that are very interesting and very nice." And for me the benefit is already a change in the way of thinking. And that's enough.

The tension and trade-offs between rigour and feasibility also become visible in the type of projects that local behaviour experts run. Limited capacities and difficulties to conduct experiments in dynamic choice environments lead these experts to focus on only a select number of politically and practically feasible projects. As a result, certain types of experiments and interventions tend to prevail. An example of this stems from

the local domain of waste management policy. Here, different municipalities, in collaboration with behaviour change consultancies, have run similar experiments to counteract the tendency of citizens to incorrectly place household waste in neighbourhood collection containers. Another popular 'quick win' is redesigning government letters and communications, for instance those directed towards citizens with an outstanding debt, or companies that are required to take action concerning building permits or maintenance obligations. Choice architecture of letters is appealing as it is simple to administer, scalable and easy to reproduce (John 2018). Given these characteristics, this particular intervention lends itself well for experimental evaluation. Compared to this informational form of choice architecture, physical forms are perceived to be less feasible. The swimming pool case for instance illustrates how the redesign of the building is costly and difficult to replicate. Also default changes seem less feasible as they, in the words of an interviewee, tend to be 'anchored more deeply' in policy.

To summarize, the local behavioural practice seems to have unlimited scope and potential. Cases eligible for a behavioural approach are diverse and numerous. This potential, however, appears less potent in practice: experts are not always able to work methodically in an 'incrementalist' environment full of constraints. Operating in the context of trade-offs and tensions between rigour and feasibility, they tend to limit themselves to quick wins that have already been tried and tested elsewhere. There is, however, room for the development of more innovative and complex projects. Nonetheless, the way in which experts manage to strike a careful balance between rigour and feasibility is undoubtedly a feat. It is demonstrative of their ability to adapt to their local institutional context as well as their ability to integrate such practices in a sustainable manner. That their practice is in part adjusted – in an attempt to meet environmental demands regarding speed, visibility and pragmatism – reflects their ability to align and understand competing demands.

***Politics***

The analysis will now turn to the political dimension of local behavioural expertise. The focus will not be so much on political institutions, like the city council or the board of the Mayor and Aldermen, but more on 'the political'. This entails issues surrounding work practices that (can) become controversial, stir up conflict and struggle and (may) lead to disturbances. The central focus is on the relation between behaviour experts and external audiences who may have certain expectations, desires, thoughts and/or feelings.

> One of the BIG'R members talks about the background of the swimming pool case: 'A colleague came to us and thought: "Why don't you show us

> what you've got? Seems futile, but just show me something, have a crack at it." And now we've also been asked for other projects. For instance, in relation to the area around Feyenoord City [the new soccer stadium]. That's what it's about: you've got to have a few marketable successes that people can explain to their wife and kids.'
>
> The expert from The Hague mentions her quest for legitimacy and resources. It is 'a chicken and egg situation' in her eyes. She needs good examples to get more resources; but in order to get these examples, she needs more resources. For now, her goals therefore are to realize quick wins, exchange knowledge, and build and sustain real commitment. 'So that we get more time and room to act and can really invest in this effectively'.
>
> A behaviour expert asks her: 'when you talk about commitment from policy, who do you mean exactly? The directorates? Or the town council? Because here (…) we do not focus so much on the aldermen. We try to keep it away from them actually. We really consider this to be an instrument. It's in the tool box. It's a form of management… and the more boring it sounds, the better it is…'
>
> When the meeting is over, we walk out of the city hall to the Erasmus University College, where BIG'R has organized a public debate about the ethics of behavioural insights. With the title: 'Is BIG'R Brother nudging you?' The debate is introduced by an Erasmus University professor who is connected to BIG'R. The professor explains why the public debate was arranged: 'The biggest mistake we can make is not taking the public along with us. We've got to be transparent in what we're doing and we'll have to hold that conversation in any case. Because we need support of the communities where we want to change something, if we want to do so sustainably. Remember that it's really a rather powerful team: government and science working together. That means that we should be monitored. Because it simply is a very strong alliance.

The fieldnote shows that the political dimension of behavioural expertise is mostly a matter of *legitimizing:* experts want to promote the utility and merit of their innovative way of thinking. In an abstract way, this is necessary to ensure that the aforementioned position and practices stem from a municipality wide authority. More practically, it is necessary to gain resources. The pressure that is experienced to prove oneself becomes apparent in the fieldnote through the recurring call for 'marketable successes'. This may sound like marketing language, but it essentially

refers to the importance of collecting successful field experiments. An example of such a success story is the aforementioned letter experiment run by the City of Enschede in order to stimulate the compliance of debt holders. The experiment resulted in a 30% increase in debt holders taking action. This 30% increase was promoted as the key result as part of a pronounced evidence-based narrative (see Jansen et al. 2017) which, in the words of an interviewee, helped in 'marketing' the approach. A subtle choice in this narrative was to emphasize the *relative* effect size of 30% instead of the *absolute* increase of 11%. This illustrates how behaviour experts make tactical choices in their analyses so as to make their story 'marketable' to a broader audience. They harness the persuasive power of quantified results from field experiments, which are especially persuasive when presented through the form of powerful visuals and 'killer charts' (Stevens 2011). Such results can become crucial soundbites which have the potential to shape further policy decisions – not just because of their analytical content but also due to their simplicity and elegance (also see Feitsma 2018b).

The fieldnote is also illustrative of the ethical debate on nudging (e.g. Leggett 2014), and the 'salience' of this debate in the field. In general, local behaviour experts do not appear to be confronted with much resistance of an ethical nature. They also cast doubt upon critical voices – here, the often-cited critique of citizen manipulation is largely deemed invalid. Common rebuttals are that behavioural influence is inevitable and that it would therefore be better for government to steer citizens in the 'right' direction. Furthermore, it is said that behavioural insights are employed transparently in relation to uncontroversial policy challenges (also see Jonkers & Tiemeijer et al. 2014). Some experts question whether behavioural policy should be the subject of ethical debate in the first place. The expert from the field note, for example, views his practice as a purely instrumental matter that should not be the topic of political or moral debate: 'We really consider this to be an instrument (…) and the more boring it sounds, the better it is…'

In the background, however, the ethical debate does appear to have a certain degree of influence on the thinking and actions of local behaviour experts. Even though they experience little direct resistance in their own political-administrative environment, they are nevertheless alert and inclined to have a rebuttal or counternarrative ready to defend and legitimize their practices. Thus, the mere *possibility* of resistance can already put pressure on them:

> I'm much more afraid of the framing of nudging as manipulation. (…) In his column in the NRC, Bas Heijne [an influential opinion maker] once wrote something about nudging as being manipulation. And recently I read something from somebody else in an article: "I do not

> wish to be nudged by the government." That got me scared. Then obviously it's placed in a completely different context. Nudging then comes to get co-opted and framed negatively by people that apparently don't realize that they are being nudged in the supermarket on a daily basis. If you do something, you're being influenced, but by not doing something your equally getting influenced. As long as you preserve freedom of choice, it's not manipulation if you ask me. But the point is that you don't want to get in that defensive role in the first place. It doesn't help. It's so easy to frame it negatively. That's something that you've got to bear in mind. That's something which we're also struggling with.

Ideological tensions do, however, come to the fore sometimes. An interviewee mentioned that he had experienced resistance with regards to policy that was aimed at reducing poverty. Plans to experiment with deregulating nudge-interventions that were in the best interests of beneficiaries were not supported by a predominantly right-wing coalition. Instead, the coalition stressed the responsibility of citizens to take measures for themselves. Another example of a 'politically sensitive' behavioural intervention that emerged was the 'youth ATM': a project where young people could ask questions about their situation on a kind of ATM, and accordingly via a printed-out money bill were informed about how much money they missed out on yearly, what they could do with this money, and where they could find more financial help. This campaign was met with resistance because of its name 'free money', devised in collaboration with the target group and partners, which evoked the idea of the City as a cash machine. Eventually a different name was chosen and the campaign went ahead. Both of these examples demonstrate that the political resistance is not so much about the instrumental nature of behavioural policies, but rather about the normative ends for which they are employed (also see Tannenbaum et al. 2017).

It is also the case that local behavioural experts will explicate the ethical debate themselves. This also emerges in the fieldnote, in which a public discussion is organized by BIG'R. The intention of BIG'R to continue to engage and deliberate with citizens about this – 'The biggest mistake we can make is not taking the public along with us.' – suggests that local behaviour experts realize that they do not operate in an apolitical space and must instead create space for public deliberation. The organizing of a debate is also seen as an opportunity to garner both explicit and sustainable support.

As a whole, the politics of local behavioural expertise mostly seems to be a matter of proving one's own effectiveness. For the field this is a comfortable form of politics, as

it fits seamlessly within both the discourse of evidence-based policy and the ambition to strengthen the 'strong alliance' between science and policy. Ideological tensions are rare and seem to operate more in the background. Local behaviour experts are nonetheless aware of (potential) conflict, and they react to this by preparing for and pre-empting the debate.

## 4.6 Conclusion and discussion

### *Concluding findings*

In this chapter, the recent emergence of behaviour experts in local Dutch governance has been explored by means of three dimensions – position, practices and politics. Table 4.2 shortly summarizes the findings per dimension and considers what this means in terms of actual opportunities for and limits to local behavioural expertise.

**Table 4.2: Opportunities for and limits to local behavioural expertise**

| *Dimension* | *Position* | *Practices* | *Politics* |
|---|---|---|---|
| *Empirical patterns* | - Explorative stage<br>- Limited capacities<br>- Predominant decentralized organization<br>- Dilemma 'traceability versus ownership' | - Richness of behavioural challenges<br>- Strong rationalist ambition<br>- Softer practice<br>- Dilemma rigour versus feasibility | - Focus on proving own effectiveness<br>- Little directly experienced ideological resistance<br>- Nonetheless emphasis on ethical reflection and deliberation |
| *Opportunities and limits* | - Modest potential due to explorative stage and limited resources<br>- Field nonetheless performs and develops further despite limited capacity | - Expertise is applied broadly<br>- Partial unused potential due to satisficing and focus on quick wins<br>- Field strategically copes with competing institutional demands | - Free rein within a pragmatic political-administrative culture<br>- Field deals consciously and strategically with (possible) resistance |

### *Contributions to behavioural policy from different levels*

The insights emerging from this study evoke a mixed picture of the opportunities for behavioural expertise on the local level, compared to the central state level. On the one hand, practices seem to have less potential given the relatively limited capacities that exist at the local level. The local behavioural landscape currently consists of

informal units with little operational and financial capacity and resources. This is seemingly at the expense of the potential for rigorous behavioural practice. Chaotic and dynamic policy arenas do not always lend themselves to the aspirational level of scientific rigour that is associated with behavioural expertise. Behaviour experts must navigate this context pragmatically by considering the trade-offs that exist between rigour, innovation, and effectiveness. The risk is that they only dare to engage in projects that deliver 'quick wins', reducing the policy diffusion that takes place to a more superficial process of *imitation* (Shipan & Volden 2012). The risk here is that part of the potential of behavioural expertise remains unused. This point mirrors a broader public administrative debate about whether municipal governments have enough *administrative power*, i.e. possess the needed resources and qualities to carry out their tasks successfully (Boogers & Schaap 2007). Compared to central state institutions, municipalities typically operate with fewer financial resources and less operating capacity. Both of these factors greatly reduce the opportunities and freedom to engage in strategic innovations. Although external collaborations with behaviour change consultancies and universities can partly compensate for this reduced capacity, this financial reality ultimately shapes the potential of local behavioural practice.

However, it could also be argued that behavioural insights are especially well-suited to applications on the local level. To begin with, local behaviour experts are already actively engaged with a multitude of behavioural challenges. This broad and rich area of application is partially due to the fact that local governance is largely concerned with a range of practical and visible matters, like keeping the city clean, safe and accessible (Barber 2013). Dealing with such, as an observed expert called it, 'practical matters', provides an abundance of behavioural challenges in practice. Moreover, local behaviour experts can get started with those challenges relatively easily given their proximity to citizens, and their capacity to act as 'direct choice architects' (Feitsma 2019). Lastly, the lack of directly experienced ideological resistance highlights an opportunity. This gives local experts the legitimacy and freedom to try and test their ideas in practice. This is in accordance with the idea that the local political-administrative culture tends to be more pragmatic and depoliticized (Barber 2013).

The medley of opportunities for local behavioural expertise became apparent during a workshop at the first annual conference of BIN NL ('Dag van het Gedrag'), organized in November 2017. Through the use of an online survey with 22 participants, word clouds were created concerning the perceived contributions of the different levels of government to behavioural policy. In these word clouds, the contribution of municipalities was mostly associated with their proximity to citizens. Municipalities are in direct 'contact with citizens' and have 'practical experience',

'local knowledge', and 'perceptiveness'. The contribution of ministries on the other hand was linked to their capacities and broader strategic function. Ministries have 'budget', 'brainpower', 'operational expertise' and 'strategic space' to collect, develop, spread and anchor knowledge. This brainstorm exercise is illustrative of the diverging ways that different levels of government can contribute to behavioural policy.

***Making progress: how?***

How can local behaviour experts progress in their journeys of institutionalization? I end with three recommendations, which are all oriented around the concept of 'adjustment'. Making progress is seemingly dependent on the extent to which a complementary *adjustment between different levels of government* is realized. This refers to a greater understanding and appreciation of how different levels of government can contribute to behavioural policy. In the context of the central state, behaviour experts may be better suited to roles which could be characterized as 'strategists' or 'networkers'. Local experts, however, may find themselves in an environment which is better suited to the experimentation of innovative techniques in practice. At present local experts seem to be active in their role as knowledge brokers, engaged in the launching of intra- and inter-organizational networks, developing relations with external partners (like behaviour change consultancies and universities), collecting successful examples, and disseminating behavioural insights (e.g. via lectures, workshops and network meetings) (also see Feitsma 2019; 2018c). This knowledge brokerage role on one hand fits with the field's explorative stage of development; on the other hand it is also up to local experts to make use of their relative proximity and develop further in the role of direct knowledge applier.

Organizing *adjustments between local behavioural practices* constitutes the second recommendation for further institutionalization. Currently, the field is developing in a multifaceted way – manifestations of this can be viewed in terms of positioning (e.g. organized as specialized unit or loose network) and practices (e.g. based on rigorous science or more on 'intuition'). In a sense, this developmental trajectory mirrors a form of *experimentalist governance* (Sabel & Zeitlin 2012) in which governments together seek optimal arrangements by initially developing a diversity of arrangements and accordingly evaluating and adjusting them (Scherpenisse et al. 2016). This quest was, in the words of the Utrecht-based behaviour expert, phrased as the 'dancing with behaviour'. This 'dance' was exemplified in the intermunicipal G6 gathering, where extensive reflection was given on the different approaches. The main focus was on sharing experiences, allowing experts to both gain awareness of the gamut of practices in existence and to adjust their own practice if needed. This made possible a more productive form of policy diffusion, based on a process of *mutual learning* (Shipan & Volden 2012). As the Rotterdam-based expert said: 'And if

something doesn't work? Let it then fail spectacularly so that others don't have to make the same mistake.'

The final recommendation for successful institutionalization involves the more abstract notion of *adjustment between competing policy demands*. This chapter highlights the resourcefulness and skill with which local behaviour experts find workable compromises and navigate inherent tensions between competing institutional logics. At first glance, the mantra and rationale of smart 'evidence-based' policy seems to be at the forefront of their practices. Behind this façade, however, their practice becomes more 'liquid' and more oriented towards sustainable integration within the existing institutional order. The shift to a softer use of evidence that reconciles rigour and political and practical feasibility is a good illustration of how adjustments can be made between competing policy demands (also see Feitsma 2018b).

The ability of local behaviour experts to adjust demonstrates that their work requires more than just an ability to stay adept of new insights emerging from the growing school of behavioural economics - despite the emphasis placed on the dominant policy idiom of 'policymaking using *behavioural expertise*' (Jonkers & Tiemeijer 2014; italics JF). Rather, the experts' work also demands political and administrative astuteness in order to be able to operate tactfully within a context of constraints, obstacles and contradictions. To capture and appreciate this dimension more fully, the foundation of behavioural expertise could more accurately be depicted as a matter of 'changing behaviour using *policymaking expertise*'.

*Chapter 5*

# Professionalizing the Behavioural State

*Governments are increasingly embracing behavioural science to improve their policies, reflecting in the rise of the so-called 'Behavioural Insights' movement. Based on half-open interviews with these upcoming Behavioural Insights experts in Dutch government, this chapter describes how they are developing their occupation, as a possible attempt to professionalize. The key issue for us is whether behaviour experts are uniform or fragmented. Our findings reveal that they are both. On the one hand, behaviour experts are unified in that they all apply behavioural science to policymaking, and to some extent rely on similar theories, methods, tools, and knowledge platforms. However, beyond these basic uniformities, behaviour experts are widely fragmented. As they are still at an explorative phase of development, possess high degrees of autonomy, and draw from a broad body of knowledge, we expect behaviour experts to remain a fragmented group in the near future.*

---

Chapter 5 is co-authored by Joram Feitsma and Thomas Schillemans. It is published as a chapter in Edward Elgar Handbook of Behavioural Change and Public Policy. Joram Feitsma has designed the study, collected and analysed the data, and drafted the manuscript. Thomas Schillemans has contributed by helping specify the research questions and theoretical framework and by critically reviewing and editing the manuscript.

Feitsma, J.N.P. & Schillemans, T. (2019). Behaviour experts in government: from newcomers to professionals? In H. Strassheim & S. Beck (Eds.), *Handbook of Behavioural Change and Public Policy.* Edward Elgar Publishing.

## 5.1 Introduction

Bureaucrats have always been in the business of behaviour change. Nevertheless, now that behavioural science has become more popular through bestsellers like *Nudge* (Thaler and Sunstein 2008), that business is becoming more and more elaborate, explicit, and scientific. Somewhat of a 'behavioural turn' has been invoked within a number of governments in the last decade with bureaucrats turning to insights from the behavioural sciences, behavioural economics in particular, to upgrade their policies. Put briefly, these 'behavioural insights' challenge the still dominant model of *homo economicus*: the self-interested utility maximizer with unlimited information and self-discipline. Instead, a different model of human rationality is put forward, which is fundamentally bounded (Simon 1957) and shaped by fast decision-making heuristics (Tversky and Kahneman 1974). When governments acknowledge and tap into the bounded rationality of humans, their policies are likely to become more effective, efficient, and better underpinned.

There are many different ways to study the behavioural turn. Scholars have studied, *inter alia*, the effects (e.g. Sunstein 2014), ethics (e.g. Bovens 2008), and legal aspects (e.g. Alemanno and Sibony 2015) of behavioural policies. This chapter, however, takes a different angle. It starts from the observation that the recent behavioural turn has not only introduced new ideas (e.g. heuristics and biases), new language (e.g. 'choice architecture') and new tools (e.g. 'nudges'); it has also changed the occupational landscape of government. The chapter studies these changes, and explores the kinds of practices that have emerged to institutionalize behavioural insights (see Lunn 2014; John et al. 2009; World Bank 2015). It examines the behavioural turn in terms of a new profession entering the policy process: the 'Behavioural Insights Teams', 'Behaviour Changers', 'Nudge Experts', 'Choice Architects', and so forth, which are dubbed 'behaviour experts'. Their emergence raises basic questions – who are they?, what do they do?, how are they developing? – but also deeper ones – what do they add?, how established are they?, are they becoming a fully-fledged profession? This chapter answers the basic questions and further explores the deeper questions within the context of Dutch government. It describes how Dutch behaviour experts are developing their new practice. The key issue revolves around the uniformity or diversity of this new occupation. Are behaviour experts developing into a distinctive, cohesive and exclusive profession in the policy process or is the development of this new occupation more scattered and diverse? The chapter will show that, although the profession is still in its infancy and to some extent homogeneous, it is currently more fragmented than cohesive. Behaviour experts share a small common core, but beyond that they rely on different theories, epistemologies, and tools.

## 5.2 Professionalization mechanisms

To describe how the occupation of the behaviour experts is developing, we draw upon theories of professionalism (Abbott 1988; Freidson 2001), and in particular of professionalization (Larson 1977). With professionalization we refer to the process in which occupations manage to mobilize themselves, become established, form an exclusive and cohesive identity, and protect themselves against external pressures. The further they go in this process, that is, the more powerful, viable and sustainable they get, the more they become professions. Professionalization has traditionally only been relevant for pure professionals, such as doctors, engineers, and lawyers. However, as part of a new kind of professionalism – one that is more hybrid and less institutionalized and content-focused (Noordegraaf 2007) – many other occupations have since positioned themselves on the road towards professionalization; for instance, agents (van der Meulen 2009), activation professionals (Schonewille 2015), and strategists (Noordegraaf et al. 2014).

In this chapter, we interpret the development of behaviour experts as a case of early professionalization. However, we do not claim that behaviour experts are on the road towards full professionalization comparable to the pure professions. Neither do we make the claim that they should professionalize to that degree, that the process is irrevocable, nor that they understand their own development in those terms. Our aim is not to follow up on Wilensky's (1964) question about whether we are experiencing 'the professionalization of everyone'. Rather, and more modestly, we take professionalization as a theoretical lens that is useful because it gives structure to the ways in which occupations establish themselves, organize themselves internally and liaise externally. It allows us to describe and understand how the field of behavioural expertise evolves.

Professionalization is concerned with how occupations arise and manage to persevere. There are particular professionalization mechanisms which are at the bottom of this process. Building upon the categorization of Noordegraaf et al. (2014; see Thorstendahl 1990), we identify four of those mechanisms: knowledge, standards, positions, and identities. The knowledge mechanism refers to how professionals regulate their work through building cognition (Thorstendahl 1990), content (Noordegraaf 2007), or a body of knowledge and skills (Freidson 2001). This mechanism concerns how they build and demarcate an exclusive body of knowledge, supported by specific schools, training courses, professional journals and knowledge platforms. The standards mechanism refers to how professionals regulate their work through creating and regulating standards about competency, ethics, quality and instrumentation (Thorstendahl 1990). The positions mechanism refers to how

professionals regulate their work on a political level (Faulconbridge and Muzio 2012) through claiming autonomy and legitimacy (Freidson 2001). It involves professionals developing strategies to deepen their links to the organization and justify their exclusive position. The identities mechanism refers to how professionals regulate their work through building a shared social identity (Thorstendahl 1990). This mechanism addresses how professionals perceive their own roles in relation to the wider context. It points to the particular heroes, model practitioners, and classic reads that help to define a profession's collective identity.

These mechanisms point at both an internal logic, geared towards finding internal closure and coherence, and an external logic, geared towards positioning oneself vis-à-vis others and other occupations. To professionalize, according to our framework, means to build and legitimize occupational boundaries, within which one develops one's own knowledge, standards, positions, and identities. The pure professions have advanced these mechanisms most consistently and elaborately, yet any existing occupation needs at least some level of shared knowledge, standards, positions, and identities. Together, these four mechanisms form the analytical lens – displayed in Table 5.1 (derived from Noordegraaf et al. 2014; Freidson 2001) – with which we observe salient aspects in the upcoming field of the behaviour experts.

**Table 5.1: Four professionalization mechanisms**

| *Knowledge* | *Standards* | *Positions* | *Identities* |
|---|---|---|---|
| - Education and training<br>- Body of knowledge<br>- Knowledge platforms | - Required competencies<br>- Quality standards<br>- Ethical codes<br>- Standardized instruments | - Legitimacy<br>- Autonomy | - Heroes and role models<br>- Role identities |

## 5.3 Methods

Our study on the development of behaviour experts is based on desk research, a series of ten preliminary interviews (with behaviour experts and academic experts), and then a series of 24 half-open interviews with 35 behaviour experts in Dutch government over the course of 11 months (April 2015 until February 2016). Interviewing behaviour experts helped us to acquire 'thick descriptions' (Geertz 1973) of their work and professionalization activities. During the half-open interviews, we were guided by a set of sensitizing topics (including professional background, organization, practices, successes, challenges, and developmental processes). The interviews were recorded, selectively transcribed and turned into individual field

reports, structured along our professionalization framework (i.e. knowledge, standards, positions, and identities).

In selecting our respondents, we looked for self-proclaimed behaviour experts who use behavioural science on a structural and explicit basis. The latter requirement of explicitness means that behavioural science directly informs our interviewees' practices and not just implicitly or in retrospect – which Lourenço et. al. (2016) describe as 'behaviourally-informed' and not 'behaviourally-aligned'. To map the presence of these behaviour experts in Dutch government, we started with desk research. A few studies were useful here (e.g. Dorren 2015). We also scanned online content (e.g. reports, LinkedIn profiles, websites) searching for a broad range of relevant keywords such as 'nudge', 'Behavioural Insights Teams', and 'applying behavioural insights'. Being flexible about these terminologies was important as the jargon of behaviour experts tends to differ and change, even though their practices are alike. Besides doing desk research, we asked respondents about their knowledge of other behaviour experts in the field. This snowballing technique provided us with more relevant sources than simple searches could accomplish, as many pioneering behaviour experts do not have clear public profiles. We also kept a list of active behavioural units that we verified with both our respondents and an online community within government that shared knowledge about the use of behavioural insights.

All of the behaviour experts we interviewed work in Dutch central government. We found most BI-related activity in this locus, with behaviour experts coming from nearly all Dutch ministries and many aligned agencies. The 35 behaviour experts we interviewed are already part of twenty different governmental organizations. Our locus of research contrasts with most studies on behavioural policymaking, which tend to focus on the frontrunners in the field: resourceful units, often part of an Anglo-Saxon policy culture, like BIT UK (e.g. John 2014). However, beyond these frontrunners there is a broader field of emerging behavioural practices 'all over the world' (Whitehead et al. 2014). If we are interested in how this field is developing, in all of its variety, it is important to study these less outspoken cases as well. The Dutch government is such an underexplored case (for exceptions, see Lourenço et al. 2016; Stinesen and Renes 2014), featuring a broad palette of explorative behavioural policy practices.

### 5.4 The rise of the Dutch behaviour expert

At face value, behavioural expertise seems to have sparked the interest of the Dutch government. More and more initiatives are being launched to apply behavioural

expertise to the policy process. These are emerging across agencies, at different stages of policymaking, as the 35 interviewed behaviour experts work both in central government (ministerial departments) as well as in agencies further removed from the centre (responsible for knowledge distribution, implementation, regulation, and enforcement).

Most behavioural practices have been established in recent years, rendering behavioural expertise still in an explorative and dynamic phase of development. The publication of Thaler and Sunstein's *Nudge* in 2008 was an important take-off point. When 'Nudge' entered the Dutch policy discourse, officials were increasingly being informed about the value of behavioural expertise for policymaking. This learning process has been fuelled by reports of the influential scientific council for government policy (e.g. Jonkers & Tiemeijer 2014). These reports were followed by a string of reports from other official think tanks and advisory bodies, culminating in a memorandum to Parliament (Ministry of Economic Affairs 2014). About thirty interviewees have started their practice only after 'Nudge' was introduced in Dutch government, somewhere between 2009 and 2015. During that period, seven interviewees were specifically hired by agencies because of their behavioural scientific background, which confirms the increased demand for behavioural expertise at that time. Only six interviewees used behavioural insights before 'Nudge', about ten to fifteen years ago, when behavioural science was already blossoming but remained largely outside the awareness of policy-makers.

Most behavioural policy practices have not been deeply institutionalized. They are organized ad hoc, bottom-up, with limited resources and feeble links to existing policymaking processes and institutional actors. Their development depends on the ambition of individual enthusiasts and efforts of low-key group collaborations, using behavioural science only sporadically or in spare time. Therefore, the positions of behaviour experts inside their organization tend to be fragile. They need to 'survive' in their organizations, challenged by cutbacks, reorganizations, disbelievers, and the burden of having to prove their added value.

The most typical organizational role model of behavioural policymaking is the Behavioural Insights Team (BIT): a specialized government unit that applies behavioural insights fulltime, inspired by the original BIT UK which was launched in 2010 in British government. At the point of writing and researching, based on a snapshot of the field from early 2016, we discovered five BITs in Dutch central government. These are the 'Team Gedragsverandering' [Team Behaviour Change] in the Dutch Tax and Customs Administration, a BIT in the Ministry of Economic Affairs and Climate Policy, a BIT in the Ministry of Infrastructure and Water Management, a

BIT in the Dutch Healthcare Authority (NZa) and a BIT in The Netherlands Authority for Consumers and Markets (ACM). These BITs are all relatively small, with up to five members, and new, being founded after 'Nudge' entered Dutch policy discourse. The 'Team Gedragsverandering', which started in 2009, is the oldest Dutch BIT; the others were all founded after 2012.

At the same time, there are many other behavioural policy initiatives beyond the spotlights of BITs. Only focusing on BITs would be to fall for a 'Behavioural Insights Team Bias' (Sunstein 2014). Avoiding this bias is critical in the Dutch case too, as the great majority of our interviewees do not work in settings that fit the BIT prototype. Twenty interviewees operate in other types of collaborations that tend to be 'looser' and less intensive, such as informal networks, knowledge exchange groups, working groups and strategic projects. Also, seven interviewees work primarily individually (e.g. as 'Gedragsbeinvloeders' [Behaviour Influencers]).

## 5.5 Professional developments in the Behavioural Insights field

***Knowledge***

We start our analysis by turning to the knowledge mechanism of professionalization. This mechanism refers to how behaviour experts develop and regulate their own knowledge content. We will shed light on the educational background of behaviour experts, the nature of their knowledge, and the presence of overarching knowledge platforms.

*Education and training*

The professional background of behaviour experts in Dutch central government is mixed. While there is no typical schooling required to become a behaviour expert, some types of backgrounds seem more fitting than others. Lengthy behavioural scientific schooling programmes seem useful. Of our 35 interviewees, 17 people had such an official background, with academic degrees in disciplines such as social psychology, behavioural economics, communication sciences and criminology. Interestingly, six interviewees followed the same social psychology master in 'Behaviour Change' at Radboud University Nijmegen. These professionally educated behaviour experts mostly work for the formalized BITs. For instance, the BIT at the Ministry of Economic Affairs and Climate Policy consists of three members that all have a doctorate in social science (experimental economics, econometrics and social psychology). Furthermore, the members of these BITs tend to have similar professional backgrounds. For example, the 'Team Behaviour Change' at the Dutch Tax and Customs Administration includes four social psychologists, and the Dutch

Healthcare Authority consists in a collaboration between five behaviour experts, all specialized in criminology.

Yet, those who engage in behavioural policymaking do, at present, for a substantial part (half of our interviewees) not possess formal degrees in behavioural studies. Some interviewees have backgrounds that still loosely relate to their current practice, for instance having studied political science, law or public health sciences. But for others, such as those who studied chemistry, human geography or chemical technology, the gap between their original background and their behavioural practice is much wider. They became behaviour experts in less straightforward ways, like in the case of the following interviewee, who, when asked about his background, states: 'Dutch literature scholar – and by accident I then spent the rest of my life trying to get people to change their behaviour.' The threshold to become a behaviour expert, in terms of one's academic background and level of expertise, is thus relatively low at present. While one must possess basic behavioural expertise, one doesn't have to be academically trained in one of the behavioural sciences. To acquire the knowledge needed, there are other ways than formal academic training, for example, self-study, on-the-job learning and post-secondary education. Some interviewees also note that they have developed their behavioural expertise more intuitively.

*Body of knowledge*

Behaviour experts tend to draw on a shared set of foundational texts and intellectual heroes. We identified a few books, articles and reports that are particularly popular amongst the interviewees. These were Thaler and Sunstein's (2008) *Nudge*, Cialdini's (1984) *Influence*, Tversky and Kahneman's (1974) theories about heuristics and biases in decision-making, Kahneman's (2011) classification of *System I* (automatic) versus *System II* (deliberative) thinking, and models like 'MINDSPACE' (Dolan et al. 2010) and 'EAST' (Service et al. 2014). Reports from the Dutch Scientific Council for Government Policy (WRR) (and also some other Dutch policy advisory bodies) are regarded to be important as well (Dorren 2015). Next to this, behaviour experts also actively search for role models, guidelines and best practices. BIT UK is such a role model, having published about its theoretical frameworks, methods and concrete projects. Its Randomized Controlled Trials (RCTs) to test new policy measures have inspired several interviewees to conduct similar experiments. For example, the Dutch Tax and Customs Administration's BIT ran various letter experiments, testing the effects of slight changes in the wordings of tax return letters, which were highly similar to previous trials by BIT UK.

Beyond these recurring core ideas and texts, however, interviewees individually draw upon a mishmash of other theories and models. They tend to operate with a

'cloud' of theories in mind and look beyond particular scientific domains. They stipulate the importance of diversity in their accumulated 'body of knowledge', like this interviewee does: 'I mainly use social-psychological literature, but sometimes also political psychology, you know, it comes from all sorts of nooks and crannies, and demarcation doesn't seem necessary to me.' Another interviewee, similarly, notes how he draws upon various streams of knowledge in his practice, from behavioural economics to social psychology to neurology. At the same, some interviewees cherish specific models. One respondent, for example, sticks with a model from Poiesz (1999) that explains behaviour change as a three-sided matter of motivating people, strengthening their abilities, and empowering them. These kind of models help them to simplify and give structure to the often complex behaviour change cases that lie before them.

*Knowledge platforms*

While some particular theories seem more popular amongst behaviour experts and closer to the core of the behavioural practice, no specific knowledge content appears to be required in the field. This is not a great surprise given that the field knows no formal knowledge authorities. We identified neither overarching platforms nor journals that represented the particular field of behaviour experts. Behavioural expertise seems to be built in a splintered manner, with behaviour experts drawing upon their own preferred sets of theories, based on their own unique background, without much contact with experts in other agencies or regulation from external authorities that exercise control over what they (should) know.

At the same time, there are modest attempts to streamline the behaviour expert's collectively shared content. Some organizations have written public reports on how to apply behavioural science. Examples are a booklet called 'Gereedschap voor Gedragsverandering' [Tools for Behaviour Change], which translates behavioural insights into practical tools for communication campaign specialists, and a brochure 'Rare jongen, die Homo Sapiens!' [Weird guy, this Homo Sapiens!], which links these insights to the work context of social affairs. Furthermore, some organizations have conjoined in platforms to exchange knowledge, like the Behavioural Insights Network Netherlands (BIN NL), in which delegates from nearly all ministerial departments participate. BIN NL organizes general introductions, hosts internal meetings to share knowledge, and facilitates a digital library and community of behaviour experts. A similar inter-organizational platform is the programme 'Handhaving en gedrag' [Enforcement and behaviour], in which behaviour experts from different regulatory agencies sit together to build and exchange knowledge, for instance during a yearly symposium. Thus, between the different Dutch governmental organizations there is a limited number of non-directive attempts to coordinate behavioural expertise. This

is often the case in a governance system in which many organizations are relatively independent (Schillemans 2012). Within organizations, however, much greater initiative in building behavioural expertise exists. Disseminating behavioural knowledge internally (through educational programmes, talks, articles, advice, etc.) tends to be a core business of the interviewees.

*Conclusion: a cohesive knowledge core with divergent peripheries*

Behaviour experts in Dutch government draw on a cohesive core of ideas, books, reports, and role models. Beyond this shared core, the knowledge base, expertise, and ideas are broad and diverse, and to some extent uncoupled. New knowledge and experience is mostly built and shared within organizations. There are a few platforms to exchange and streamline knowledge, yet they tend to be optional and organized in a bottom-up fashion. There is also little control from overarching bodies over the knowledge that behaviour experts (should) use. Behaviour experts seem little restrained by requirements in terms of background, experience, and knowledge. Overall, the degree of occupational closure in terms of knowledge is limited. From a strict professionalization perspective, this could be seen as disadvantageous. However, some of our interviewees highlight the pros in this, as it enables them to pragmatically develop their own mixed set of skills and specialisms. Furthermore, as most behaviour experts still draw on a small shared core of ideas and texts, this still leads to some de facto alignment and convergence. As one interviewee stated: 'The surprise is that we're all working in different contexts, and have been developing our knowledge and expertise fairly separate from each other from the start, but still we're all arriving at a sort of similar approach.'

***Standards***

This section addresses the standards mechanism of professionalization. We particularly look at how behaviour experts standardize required competencies, quality standards, ethical codes and instruments.

*Required competencies*

The previous observation that half of the interviewees are not behaviourally schooled, clearly suggests that there are no formal requirements for behaviour experts. On an informal level, however, looking at the competencies deemed necessary in the field, it becomes apparent that almost all interviewees emphasize the sophisticated nature of their practice. It is 'not a gimmick, but a craft'. It is 'hard work' that requires advanced competencies to understand and control behaviour in particular areas. It appears that behaviour experts rely on two different sets of competencies. The first set underlines their scientific competencies in applying behavioural science. Behaviour experts must know what drives human decision-making, and how that

knowledge can be translated and made applicable for policy-purposes. They must also have practical experience with doing research (e.g. regarding literature study, desk research, interviewing, observation, and field experiments). The second set emphasizes the political and administrative competencies. That includes administrative experience, organizational sensitivity, know-how of policy processes and change and project management skills. Behaviour experts ideally are savvy 'change managers' who know how to build networks and gather support from key actors. This suggests that interviewees believe that they should be able to combine the best from two worlds: the rigor of behavioural science with the savviness of public policy work.

*Quality standards*

Shared quality standards are crucial for pure professionals. In that sense, the level of professionalization of behaviour experts seems to be in its infancy because our interviewees subscribe to strongly different views on what guarantees the quality of their practice. They especially seem divided in terms of what kind of knowledge counts as valid evidence. The division is between those who favour soft professional knowledge versus those who favour hard experimental knowledge, i.e. those who favour a *behaviourally-informed* approach to policymaking versus those who favour a *behaviourally-tested* approach (Lourenço et al. 2016). Or, in Aristotelian terms: those who favour a *phronesis* type of knowledge versus those who favour *episteme* types of knowledge (Parsons 2002).

The group of behaviour experts that favour experimental knowledge emphasize the evidence-based nature of their practice, and the need of 'experimentally demonstrating that it works'. Following their modern-empiricist scientific role model, BIT UK, they advocate that policy interventions should be *ex ante* evaluated through small-scale controlled scientific experiments. Doing these kind of trials, like RCTs and A/B testing with websites, is the most reliable way to predict human responses to policies, as it helps to isolate the actual effects of single interventions

> What you mainly want with behaviour change is examining and demonstrating that it works… And because we've got this increase in data, there are much more behaviours… that we can map out. So, then you can start running experiments and effect measurements with A/B testing. And then you're able to see: this intervention works better than that one.

The group of other experts, favouring professional knowledge, however dismiss RCTs as 'the holy grail'. They note that RCTs are relatively expensive, time-

consuming, and also difficult to set up (challenges involve formulating quantifiable outcomes, keeping control over the treatment groups in a dynamic field setting, and isolating the effects). On top of that, RCTs can introduce ethical questions and political difficulties (e.g. institutional resistance to effect measurement out of fear for potential negative outcomes). Experts in this group thus rely on a different, in their eyes more feasible, source of knowledge: their own professional judgment, based on a mix of common sense, logical reasoning, creative thinking, theoretical knowledge, and field experience. These behaviour experts deviate from the scientific model, stressing that 'this here is not a university' and that their job is 'not an exact science, but an iterative model' that involves making 'educated guesses' and learning through trial-and-error. They often have pragmatic reasons for their approach: limits to organizational resources force them to make do with what they have, and generally this means trading the rigor of experimental knowledge for the relevance of professional 'indications'. They find themselves in a balancing act, having to make trade-offs between certainty, rigor, and thoroughness on one hand, and feasibility, speed, and efficiency on the other. They have to make compromises: gathering just enough certainty to make valid assessments without needlessly suspending the flow of work. A result of this balancing act is that these behaviour experts must cope with uncertainty. They lack hard experimental evidence to prove their effectiveness. 'It's not rocket science', one interviewee claims, 'that's the difficulty in behaviour change, you can't promise that it's going to work'.

Even though behaviour experts de facto adhere to different quality standards, they are nonetheless united in their support of evidence-based policy as an ideal. Evidence-based policymaking thus serves as an informal shared frame of reference (or loose standard) for the behaviour experts.

*Ethical codes*

A particular area of interest for standardization is the regulation of ethical norms, given the extensive ethical debates about behavioural policymaking in the administrative (e.g. House of Lords Science and Technology Select Committee 2011) as well as in the academic world (e.g. Bovens 2008). Ethical objections, *inter alia*, state that nudging lacks transparency, doesn't promote what people themselves want and diminishes their autonomy through bypassing their reflective decision-making (e.g. Hansen and Jespersen 2013). Although these concerns have traditionally been raised in the political context of the United States, where state intervention meets relatively strong political resistance, they seem to have been transferred to interviewees' European context too. The popularization of 'Nudge' in the Dutch policy sector has not only introduced potentially innovating ways of policymaking, but also sensitive underlying ethical issues (e.g. Jonkers & Tiemeijer 2014; Van Staveren et al. 2014).

Hitherto, ethical debates have not led to shared ethical standards for behavioural policymaking. Instead, interviewees deal with ethical issues on the basis of their own moral compasses. Some disregard the ethical discussion and claim that ethics is not their area of expertise. Others reject the ethical concerns that nudging would be inherently 'dirty' or 'wrong'. They emphasize that behaviourally-inspired interventions are by their nature not more or less legitimate than other policy tools. Some interviewees do give more credit to ethical concerns, particularly in relation to the questioned legitimacy of targeting unconscious behaviour. One way to overcome these concerns is by setting prior conditions for legitimate practice. Transparency seems to be such an important condition. It has led the Dutch Tax and Customs Administration's BIT to publish openly about its existence, working methods, and nudges in a major Dutch newspaper (De Jong & Rusman 2 March 2015).

*Standardized instruments*

The one area where we see clear examples of standardization is in the area of instrumentation. Although most behaviour experts are still setting up shop, we identified some standardized techniques and methods that are popular in the field. These are mostly policy analytical tools. 'CASI 3.0' is such an example, a tool which helps communication specialists to develop a behaviourally-informed communication campaign strategy. Similar tools are the 'Gedragstoets 2.0' [Behaviour Test 2.0] and the model 'EAST' (which exists both as a general guideline and also as a deck of inspiration cards) developed by BIT UK (Service et al. 2014). What unites these tools is that they all act as action plans for designing behaviourally-informed policy and communication strategies. In several steps, in series of questions, these tools help to explicate and examine underlying behavioural assumptions in policy and campaign strategies. Underneath, these diagnostic tools subscribe to a *rational-comprehensive* (Lindblom 1959) perspective on policymaking in which policy problems are (and can be) solved through rational and systematic analysis. To illustrate, one interviewee developed his own systematic 'action plan for behaviour change' based on six steps, moving from 'describe the problematic behaviour', to 'describe the desired behaviour', to 'analyse the behaviours', to 'select the behavioural measures that match the analysed situation', to 'implement the measures', to 'test the measures with pre- and posttests and control situation' (Gemeente Schoon 2014).

Behaviour experts have also to some extent standardized the use of nudges. Nudging involves the translation of behavioural science into minor readjustments – 'psychological gimmicks' – of people's daily surroundings. Due to their small and informal nature, nudges appear relatively easy to standardize. To illustrate, one interviewee mentioned having some 'basic' nudges that she applies throughout her

work. A nudge that she standardly suggests to communication colleagues is that when penalizing companies they should always immediately offer a future solution too. Such an action perspective encourages behaviour change, she claims. It is a quick win that cannot go wrong. Standardization can however also be more challenging, especially when behaviour experts deal with complex behaviours in dynamic choice environments.

*Conclusion: some shared instruments but little standards beyond that*

Our research suggests that there are some uniform standards but very little beyond that. There are most notably no formal qualifications needed to become a behaviour expert and there are no clear ethical or quality guidelines. Organizations mostly set their own professional standards. The only area where we see more advanced standardization is in the area of instrumentation. At the same time, some informally shared standards seem to exist, for instance in relation to required competencies and quality of work. Another less formalized but nonetheless important feature of cohesion is the diagnostic nature of the job. Most behaviour experts adhere to a rational-analytical view of policymaking where they are all, some more outspoken than others, behavioural science appliers to policies.

***Positions***

This section deals with the positions mechanism of professionalization, which concerns how behaviour experts position themselves vis-à-vis others. We will mainly address their search for legitimacy and autonomy.

*Legitimacy*

The search for organizational legitimacy is a particularly thorny issue for the interviewees as they need to carve out a distinctive position in well-established policy-arenas that feature various and competing values and actors. As a result, making themselves visible, giving pitches and internal consultations and building networks, consumes a substantial amount of their time and energy. Many interviewees are, as one put it, 'struggling' to win the support from others. An interviewee states: 'Most of our time is lost by needing to convince our internal clients to do things different externally.' Part of the struggle lies in general challenges faced by any innovation (e.g. lack of time and resources, start-up problems, not being prioritized, etc.) and in the fact that behaviour experts – and their associated theories and working methods – are not always understood by their colleagues with different backgrounds. Underneath that, a more fundamental mismatch between the experts and the wider policymaking culture seems to exist. Their introduction of a new school of knowledge (behavioural economics) and increased attention for policy experiments and in-depth policy analysis, not only challenges the prevailing

knowledge paradigms (law and economics) in Dutch government, it also slows down the policy cycle and fuels it with uncertainty. This doesn't cohere well with a speedy and solution-oriented policy culture with little space for uncertainty and 'second order' activities like analysis and evaluation (Lindblom 1959). At the same time, the fact that behaviour experts are struggling doesn't wholly prevent them from making progress in finding a legitimate role, generating visibility, and overcoming resistance. These two BIT members, for example, note that their position has gained strength with the years:

> Our direct colleagues come to us quite a lot… I think in that respect we're further developed than many other regulation agencies who are still very much concerned with putting themselves on the map internally, building support in order to get colleagues to work with behavioural insights, or getting at the table somewhere at all… Getting involved in things… Well, that's something we really don't have to do. Over here, colleagues know what our role is, what our contribution is, and what can be expected of us.

*Autonomy*

As they strive towards more control over their practice, autonomy is another major theme for the behaviour experts. The formation of BITs and other specialized behavioural functions may be of importance in getting more occupational control, as it gives concrete institutional shape to behavioural expertise. These BITs do not per se need formal legitimacy to be able to exercise autonomy. One BIT was even purposefully organized unofficially and kept 'under the radar' of the organization, with one of its members noting: 'As a BIT we're not an officially founded team… rather we did it on our own initiative. With the goal to keep as much freedom in what we do and not having to operate within the boundaries of the organization.' At the same time, the autonomy that comes with not being institutionalized can be compensated by other limitations. Informal BITs have a harder time in acquiring resources and gaining moral support, and they cannot fall back on formal structures, powers and rules. Hence, in the end most behaviour experts do strive for more autonomy through further formal institutionalization.

Despite their strivings for autonomy, at the same time the reality of most behaviour experts is that they are highly dependent on many other policy actors. They mostly work *with* or *for* others. These dependencies take shape in different forms. Sometimes behaviour experts play advisory roles in existing policy projects. It also happens that they step in as trainers, teaching colleagues how to integrate behavioural knowledge. Or they become project managers, building broad networks of expertise (including

universities, consultancy bureaus, trade organizations, and citizens). The BIT at the Ministry of Infrastructure and Water Management, for example, adopts a 'network-approach' which builds a flexible team of experts around each of its projects. Through collaboration, this unit seeks to acquire the necessary knowledge and experience to engage in complex behavioural policy initiatives. For many behaviour experts, an awareness of their dependency within the broader policy system is key. They need to blend into the system, and collaborate with others, in order to work effectively. As one interviewee explains, 'Behaviour change is about aligning the organization from A to Z so that it carries out the same message and makes the same movements.'

*Conclusion: shared struggles for support in the face of dependency*

As newcomers in a well-established policy system, behaviour experts face the issue of having to build legitimacy. While some have more success than others, this issue generally tends to be a struggle. The behaviour experts seek more autonomy to increase occupational control. At the same time, their dependencies on others force them to further integrate into the existing system, trying to connect to and merge with existing practices.

***Identities***

This section turns to the identities mechanism of professionalization, which refers to the ways in which behaviour experts construct and regulate their own identity. We look at their heroes, role models, and typical organizational roles.

*Heroes and role models*

The identity of the behaviour experts partly stems from the role models they follow, such as Thaler and Sunstein, Kahneman, Cialdini and BIT UK. The fact that the Dutch field also has installed 'BITs', mirroring BIT UK in name and organizational model, reflects the exemplary role of BIT UK well. Yet, many other role models are mentioned, including academic scholars, think tanks, and fellow experts. Similar to how interviewees tend to draw on their own particular 'cloud' of ideas, they also have their own particular 'cloud' of heroes and role models.

*Role identities*

In a broad sense, our interviewees share the same quest: they are all revolutionists (or phrased more modestly: innovators) who cast doubt upon the, as one put it, 'bizarre' conventional ways of policymaking. They all believe that harnessing the power of behavioural insights results in better policies. At the same time, beyond this common quest, interviewees take on different organizational roles that express different identities. Some of these roles seem more prevalent than others. One prevalent role, taken on predominantly by about half of the interviewees, is when behaviour experts

directly engage in choice architecture**.** They actually intervene in particular contexts, for instance in school canteens to stimulate healthy eating or in local neighbourhoods to avoid littering. Sometimes this direct type of choice architecture consists in being creative, offering 'fresh looks' on existing practices and loosely suggesting small interventions. Other times it consists in doing rigorously experimental research, setting up labour-intensive RCT projects.

Yet, behaviour experts are not always direct choice architects. In (at least) three other roles, their contributions towards behaviour change are more indirect. A first more indirect role consists in making policy analyses. In this role, behaviour experts purposefully play the 'advocate of the devil', trying to challenge and (if needed) correct the deeper behavioural assumptions that underlie specific policy theories. They aim for a more thorough policy analysis through dialogue, 'constantly asking difficult questions… Why-questions, like a little kid: "Yes, but why? Why? Why?"' This role resembles Socrates' role as the social gadfly in Ancient Greece. A second more indirect role consists in managing behavioural policy projects and networks. In this role, behaviour experts set aside matters of content, and instead focus on the process of behavioural policymaking and who gets involved. They 'assemble' and coordinate networks of expertise, while managing administrational and political issues. A third more indirect role consists in spreading knowledge about behaviourally-informed policymaking. In this role, behaviour experts mainly inspire other actors. They act as intermediaries, positioned between science and policy practice, ensuring that policymakers and practitioners learn about behavioural insights in a language they understand.

*Conclusion: circuitous roads on a common path*

Behaviour experts implicitly share a small common core of role models, with much fragmentation beyond that. While they are united as BI revolutionists, they also play (and combine) different organizational roles. Table 5.2 situates these (neither all-comprehensive nor mutually exclusive) four roles described above along the lines of how directly they contribute to choice architecture.

**Table 5.2: Diverging roles of behaviour experts**

| Indirect choice architecture… | | | … Direct choice architecture |
|---|---|---|---|
| *Messenger* | *Network node* | *Social gadfly* | *Choice architect* |
| Spreading knowledge | Managing networks and projects | Making policy analyses | Running RCTs and implementing designs |

## 5.6 Conclusion and discussion

In this chapter, we have studied behaviour experts in Dutch central government. More specifically, we have studied how their occupation is developing, interpreting it as a case of early professionalization. Our study shows that behaviour experts are slowly finding their way into Dutch central government, and are being attracted by many agencies across the board. The experts are still at an explorative stage, in the process of setting up shop, making themselves seen, and establishing effective practices, with minimal resources, powers, and ties to their organization. While the behaviour experts are far from identical, they do share a few core characteristics. To begin with, they all more or less rely on a number of iconic authors (e.g. Thaler and Sunstein, Kahneman and Cialdini), role models (e.g. BIT UK), advisory reports (e.g. Jonkers & Tiemeijer 2014), tools and instruments (e.g. 'C.A.S.I. 3.0'), and loosely coordinated professional networks (e.g. 'BIN NL'). Moreover, behaviour experts all tend to view policymaking as a highly rational-scientific affair. Though some are more outspoken and rigorous than others, they all apply behavioural science to policies. Nevertheless, behaviour experts form an occupation that is currently more fragmented than cohesive. They give shape to their work in different ways, based on their own experiences, beliefs, and preferences. They rely on their own, widely diverging 'clouds' of ideas and role models.

The observed fragmentation can be linked to at least five underlying forces, suggesting that behaviour experts will remain relatively fragmented in the near feature. First, government agencies in the Netherlands are traditionally highly autonomous (Schillemans 2012), which may explain part of the fragmented development of behavioural expertise. Second, behaviour experts are relatively unhampered by the kind of restrictions and obligations that the more mature occupations face. This gives them more leeway in shaping their own practice, resulting in fragmentation. Third, behaviour experts are part of government agencies with widely differing tasks as well as interactions with citizens and companies. The fragmentation may be a result of behaviour experts customizing their practice to their particular policy context. Fourth, government agencies are increasingly withdrawing from direct interaction with citizens. This new role puts pressure on the behaviour expert's most typical role as a choice architect, directly shaping the environment in which people make choices. With the current political emphasis on co-production, agencies are forced to explore different, more indirect routes towards choice architecture, strengthening further fragmentation. Fifth, behaviour experts draw from a rather broad universe of insights from various behavioural scientific disciplines (including contradictory findings and disputed theories). This broadness allows some behaviour experts to cherish their own hobby horses, and some others to merge all

sorts of insights into a particularistic hotchpotch, strengthening further fragmentation in the field.

Our study has shown salient ways in which behaviour experts in the Netherlands are professionalizing. Given their current status as a loosely coupled group of unfledged outsiders – with innovating ideas but little resources, promising methods but still small portfolios – the question remains how strong their influence on policymaking will become.

*Chapter 6*

# Tailoring the Behavioural State

*'Behavioural Insights' has emerged as an increasingly popular approach to policymaking in governments across the globe. Experts largely present a frontstage narrative of Behavioural Insights as a coherent concept but this chapter challenges such a description. It explores how efforts to develop a global Behavioural Insights community are subject to an ongoing process of policy translation. To show how this translation works, we juxtapose findings from two independent ethnographic research projects on Behavioural Insights experts – one on experts in Australian Federal government, the other on experts in Dutch local and central government. This exploratory study highlights that Behavioural Insights at one level possesses some consistencies, including a shared narrative and shared use of a family of tools and artefacts. At the same time the field is marked by contingencies, particularly with respect to the methods used. These contingencies raise puzzling questions about where the boundaries of Behavioural Insights lie and whether its presentation as a coherent whole is of more value in a discursive sense than in a practical one.*

---

Chapter 6 is co-authored by Sarah Ball and Joram Feitsma. It is based on an earlier version presented at the annual European Group of Public Administration conference in 2018 in Lausanne. It is currently under review for publication in an international academic journal. Joram Feitsma and Sarah Ball have contributed about equally to this chapter across all research stages. They have collaborated extensively while designing the study, analysing the data, and drafting and editing the manuscript. Concerning the data collection, Joram Feitsma has collected the data of Dutch behavioural practices, while Sarah Ball has gathered the data of Australian practices.

## 6.1 Introduction

In the last two decades, a new trend of 'Behavioural Insights' has entered the global policy scene (John 2018; OECD 2017; Whitehead et al. 2018). This trend aims to infuse public policy with the latest behavioural science insights, arguing that humans are not nearly as rational as presented in the neoclassical economic models. Although Behavioural Insights is commonly presented as a coherent and well-defined program, particularly by its proponents, in this article we argue for a more differentiated and contextualized understanding of it. We argue for an understanding of Behavioural Insights that is more sensitive to the dynamics of policy translation that take place when 'soft' policy ideas travel across the global policy realm. This article can be read as an exercise in mapping the 'boundaries' of Behavioural Insights, tracing key aspects of consistency but also contingency in this emerging field. We do so on the basis of extensive ethnographic fieldwork, both having spent ample time in different sites of Behavioural Insights policymaking – one in Australian Federal government, the other in Dutch government.

***The rise and reception of the Behavioural Insights movement***

Behavioural Insights (hereafter: BI) has its most direct intellectual roots in the birth of the new school of behavioural economics since the 1970s, and the popularization of this emerging body of knowledge in the broader public, policy and political discourse. One of the trend's leading role models is the United Kingdom's (UK) Behavioural Insights Team (BIT), which began as a central strategic government unit and became a specialist in the design and rigorous evaluation of behaviourally-informed policy solutions (John 2018). While many governments are still experimenting with behavioural economics (and behavioural science more generally), this trend has been rapidly extending its circle of influence, with dozens of newly created behavioural units operating at the transnational level (e.g. the 'Foresight, Behavioural Insights and Design for Policy unit' of the European Commission), central state level (e.g. the 'Behavioural Sciences Team' of Japan's Ministry of the Environment) and local state level (e.g. the 'Behavioural Insights Group Rotterdam'). Numerous non-governmental units (e.g. 'Nudge Lebanon') have been launched as well. BI has the backing of actors within policy, politics, academia, commercial and public discourse. These actors promote BI as a vital agenda to address many of the crises facing contemporary neoliberal governments – such as obesity, personal debt and climate friction (Whitehead et al. 2018). By redesigning the decision-making environment – often referred to as 'choice architecture' (Thaler & Sunstein 2008) – the promise is that these complex problems can be tackled effectively yet unobtrusively.

BI is best seen as a combined body of policy projects, procedures, instruments, organizational designs, disciplinary outlooks, theoretical ideas, evaluation methodologies and ethical orientations. Although this body is not easily specified, we would argue that there are at least four key pillars that represent BI's 'frontstage' identity, i.e. the public impression it seeks to convey (Goffman 1959). The first pillar is the use of behavioural science, in particular new behavioural economics insights into how 'supposedly irrelevant factors' impact economic decision making. BI has drawn heavily from a catalogue of systematic errors and predispositions in human judgement that influence decision-making (Kahneman 2011; Kahneman & Tversky 1984; Tversky & Kahneman 1974). The second pillar is 'nudging', which leverages these insights for the conscientious design of 'choice architecture' (Thaler and Sunstein, 2008). Nudge is an instrument intended to leverage the behavioural science noted above, and assist in the design of more effective policy. This approach is in and of itself not new but its repackaging and reframing as 'nudging' has certainly given new life to its use within governments. Nudges can be both physical and/or informational, and can include default rules, simplified messaging, social norms, the removal of friction costs, disclosure and warnings, pre-commitment strategies and reminders (Sunstein 2014; Thaler & Sunstein 2008). The popularity of nudging has led, at times, to the use of the term to capture all of the work undertaken under the BI umbrella. The UK team was known as the 'nudge unit' (Halpern 2015) and it was claimed by John (2016: 2) that '[b]ehavioural public policy is the common currency of today's decision-makers with the word nudge being used to denote this interest'. The third pillar is the use and advertisement of Randomized Controlled Trials (RCTs) (Ames & Hiscox 2016; Halpern 2015; OECD 2017; Social and Behavioural Sciences Team 2015). RCTs are promoted with the aim to minimise biased estimates of intervention effects, particularly selection bias, through the random allocation of individuals or groups of individuals to either receive the intervention (treatment) or not (control) (Glennerster & Takavarasha 2013). A well-designed RCT should be able to demonstrate a causal link between the intervention and the outcome (Haynes et al. 2012). The fourth pillar is a wider focus on evidence-based policy making and 'what works' (Halpern & Gallagher 2015). Through the combination of iterative trials and behavioural science the evidence produced by BI experts are painted in opposition to less evidence-based approaches to policy making (BIT, n.d.-b). In fact, as noted in Einfeld (2018) there seems to be an inextricable link – a 'symbiosis' – between evidence-based policy and BI discourse (also see Chapter 3). Both draw authority from instrumental scientific knowledge extensively, and both seek to further rationalize the policy process.

Despite BI's current popularity within policy circles, it has also faced considerable resistance, coming from academia, public media and across the political spectrum.

Opponents have challenged the supposed novelty, feasibility, methodological comprehensiveness and effectiveness of BI policies (Mols et al. 2015; Science and Technology Committee 2011; Whitehead et al. 2018). Moral criticism has been offered as well, holding that that the behavioural turn puts at risk core liberal democratic values like tolerance, autonomy and human dignity (Furedi 2011; Whitehead et al. 2018). Leggett (2014) goes further, noting that, as it stands, the nudge agenda may exacerbate an already changing relationship between the state and citizens, shifting the focus from the influence of social structures towards individual behaviours, raising questions about whether responsibilities for social issues are being allocated fairly. Part of the challenge of settling this debate is the lack of empirical data on BI practice and an uncertainty about how this field will develop in the future. As this trend is still relatively novel, it remains unclear how it will actually institutionalize and how it will change governmental design and implementation processes. Moreover, the behavioural turn denotes a wide-reaching and implicit process, located at the interface of state, science and society, reconfiguring the connections between those domains in complex ways that cannot yet be wholly anticipated. Such empirical uncertainties in turn pave the way for normative uncertainties about the BI agenda and the extent to which it stands in service of liberal democratic societies. As long as it remains unclear what the emerging 'behavioural state' really entails and what may become of it, differences in views about its desirability are likely to remain.

***Towards a policy translation perspective on Behavioural Insights***

This chapter tries to contribute to the BI debate by reducing some of the empirical uncertainty surrounding the trend. We seek to describe some of the ways the BI agenda has manifested, and what kind of concrete practices it is producing. While studies already exist that map the spread of BI activities (e.g. Lourenço et al. 2016; Lunn 2012; OECD 2017), it is our contention however that such mapping attempts tend to pay (too) little attention to the broad diversity of practices and beliefs that mark the BI landscape. Rather than seeking to explore and explain such diversities in detail, these studies appear to collect practices and subsume them under the umbrella of 'Behavioural Insights' without much discrimination.

This is not to say that points of difference are wholly ignored by BI scholars and experts. BIT UK (2015) developed an acronym, APPLES, to describe the way Administration, Politics, People, Location, Experimentation and Scholarship played a role in successful institutionalization of BI practice. In addition, the European Commission (Lourenço et al. 2016) and the OECD (2017) have also pointed to how practices may vary in terms of their stage of development, organizational design and general work methods. Lourenço et al. (2016) distinguish between projects that are more loosely 'behaviourally-informed' and those that are more rigorously

'behaviourally-tested'. The OECD (2017) has reported three common organizational designs; a diffuse, central steering or project-based approach. Lastly, Strassheim and Korinek (2018) have investigated the 'Many Varieties of Behavioral Public Policy', identifying four different international clusters of behavioural practices based on their organizational configurations. While insightful, we would contend that most attempts to differentiate between BI practices remain focused on organizational and generic differences. In other words, the empirical study of BI practices seems to be more oriented on what binds practices together than on what separates them.

Drawing on the diverse policy science literature on policy translation, we would argue for a closer look into what separates them – the contingent practices. We challenge the idea that there is a fixed and universal BI 'brand' trickling down homogeneously from its founders and champions. This idea reflects the concepts of used in diffusion literature which focuses on the specific enabling and obstructing conditions for the uptake of a policy. For instance, at the macro-level, structural, political or economic factors can determine whether a policies gets taken up or not. At the micro-level, the influence of individual persons and organizations, along with their resources, preferences and ambitions, has to be considered (Mintrom 1997; Nutley & Davies 2000; Rogers 1983). These processes are importantly both a technical and a social process (Rogers 1983: 4). Critics note that the diffusion approach fails to address the ways context and practice modify these innovations when they are adopted by others. Stone (2012) notes that '[d]iffusion approaches exhibit a fascination with the conditions for transfer rather than the content of new policies' (Stone 2012: 485). Moving beyond thinking about the process as instrumental adoption, researchers have also explored how innovations shift and change during the learning process. This work is commonly referred to as policy translation and focuses on how innovations and 'soft' policy ideas may transform as they travel across contexts (Freeman 2009; Ingold & Monaghan 2016; Johnson & Hagström 2005). This transformation is viewed as a form of bricolage (Freeman 2007; Stone 2017), selecting and assembling pieces of ideas that match well with the local context. A bricolage that may even combine a series of ideas into a stable 'standardized package' (Fujimura 1992). Together, these literatures cast doubt upon the idea that a megatrend like BI, however coherently portrayed at the frontstage, will maintain its professed universality and fixedness once it travels across diverging political-administrative contexts.

Moreover, it ought to be considered that BI is a rather wide-ranging concept, loosely combining four broad pillars which all already have their own history. As a result, the intellectual and institutional heritage of BI is rich and far-extending. Its core ideas (or at least the broad essence thereof) have been around for much longer (e.g. the old

school of behavioural economics school championed by Herbert Simon), and so too have been its core methods (e.g. RCTs were long introduced in the domain of medicine) and techniques (e.g. the use of 'nudges', although not explicitly phrased as such, is well-established within government). Within this context, it seems plausible that a heterogeneous BI landscape would emerge, as experts can draw from so many different elements from the rich heritage of their field.

As we will demonstrate in our chapter, under empirical scrutiny BI practices indeed appear to be subject to the dynamics of policy translation in stronger ways than thus far has been documented. While at the frontstage one can typically observe a performance that conveys a degree of coherence and consistency, at the level of backstage practice the field seems in the midst of an ongoing translation process, being tailored to local institutional contexts in contingent ways. This chapter further explores apparent consistencies as well as the meaningful areas of contingency in the BI field, particularly zooming in on methodological differences. Finally, we will argue that these areas of contingency point to important questions regarding the boundaries of BI and point to a need for greater critical scrutiny of its intent and purpose in policymaking.

**6.2 Methods**

This chapter is drawn from two independent ethnographic studies undertaken between 2014 and 2018 studying BI experts within Australia, with a focus on Federal Government, as well as experts within various Dutch central and local government agencies. This study doesn't represent a comparative study of two countries, it instead uses the teams and their authorizing institutions as its unit of analysis. This leads to a greater focus on organizational context over and above politico-administrative contexts. At times politico-administrative elements will be considered but as they were not the focus of the original studies they will only be engaged with briefly here. We do however believe further study into this would be deeply valuable. It is also important to note that these two cases are not intended to be taken as 'typical' cases. This is because it is not yet understood what a 'typical' BI unit might look like. We do acknowledge the UK as the prototypical case, but in this early, exploratory stage it is challenging to define what a 'typical' team may look like in practice. This goes to the very subject we are exploring, in considering how the concept of using behavioural insights is operationalized.

The authors believe that in order to further complement existing research there is a need to undertake interpretive qualitative research into the contemporary behavioural turn. In this light, ethnographic research presents a valuable, yet often

undervalued, perspective in public administration (Rhodes 2014). It is valuable because in the public eye and in interviews or surveys, officials and public servants may choose not to give reasons for decisions they have made, or may provide an 'acceptable' or less contentious story (Howlett et al. 2009: 7). They have a tendency 'to create second order accounts which conform to the expectations of the interviewer and especially to the implicit belief that actions must always have "reasoned reasons"' (Stevens 2011: 6). Next to this, institutions, departments and officials construct their own frontstage so as to manage their public appearances, while they may display different sets of behaviour in the backstage area – where they can prepare for but also might deviate from their frontstage narrative (Goffman 1959). An ethnographic approach allows access to the hidden workings in the backstage (Van Hulst 2008) and the policy decisions which occur 'beyond the record of formal investigation' (Howlett et al. 2009: 7).

This chapter represents only an exploratory 'conversation' between two distinct data sets. We do not intend for this research to produce representative or generalizable findings. Our chapter might be viewed as an exercise in conducting comparative hypothesis-generating case studies undertaken through ethnographic methods (Collier 1993). Our case studies serve to highlight the shared similarities but also the inconsistencies between BI practices as the 'examination of two or more cases [serves] to highlight how different they are, thus establishing a framework for interpreting how parallel processes of change are played out in different ways within each context' (Collier 1993: 108).

*Australia*

One half of the research data comes from an exploratory ethnographic case study of the Behavioural Economics Team of the Australian Government (BETA) and their colleagues and government partners. The central research question for this study explored how the idea of using behavioural insights was constituted both in narrative and as it was operationalized in practice. Time in the field was principally spent at BETA's offices. In total 47 days were spent in the field resulting in approximately 350 field hours in total. 19 formal interviews of between 30-50 minutes were also undertaken. Eight interviews were with internal staff, three were former members of BETA, and eight were members of external partner agencies. These interviews were semi-structured, and focused on gathering further insights into the formative years leading up to the establishment of the team and exploring the project selection processes. Initial participants were recruited during fieldwork using purposive non-probability sampling methods. The snowballing technique was used to recruit other potentially valuable interview subject.

*The Netherlands*
The other half of the research data is drawn from a research project that focuses on the rise of BI experts in Dutch government and their impacts on the policy process. Ethnographic fieldwork was conducted within a period of four years. The research started out with a focus on the central governmental level, interviewing 35 experts across 20 different organizations. The interviews were semi-structured and concentrated on getting a general understanding of the experts' work, asking about daily activities, routines, networks, successes and challenges. The studied sample consisted of self-proclaimed policy experts who explicitly claimed to use behavioural insights on a structural basis. These experts were identified by means of an initial document study and use of the snowballing technique. Next to interviewing, 55 hours of short-term participant-observation at different sites was done in order to grasp the day-to-day realities within BI more clearly. In addition, a four-month fulltime secondment to a ministerial behavioural unit was part of the research. Later in the research process, the focus shifted to experts in local governance, interviewing 15 local experts and doing another 19 hours of short-term participant-observation. Lastly, a focus group with various BI experts was conducted so as to validate overarching findings.

Our analysis is structured in two parts. We start by exploring consistencies, i.e. aspects of BI that appear to be maintained as they travel across local contexts. Based on our ethnographic fieldwork, we explore what is consistently captured by the concept of 'Behavioural Insights'. The second part is dedicated to uncovering contingencies in the field. These areas of contingency seem less visible as they reside at the level of backstage action, while absent at the level of frontstage talk (Brunsson 2007). Nonetheless, their meaningful implication is that the BI landscape is more heterogeneous and less coherent than one might infer from its public appearance. Importantly, the consistencies and contingencies we have identified are not so much part of a clear-cut 'road map' out there for the field to follow. Rather, their presence seems more implicit and emergent; they form part of a tacit understanding or 'hidden curriculum' of BI that is still 'in the making'.

## 6.3 Consistencies in the Behavioural Insights field

***Shared frontstage narrative***
As noted in the introduction, the literature and policy documents produced by experts appear to present at least four key pillars; behavioural science, nudge, RCTs and evidence-based policy making. Both authors witnessed the promotion of all of these in the frontstage narrative presented to other experts, policy makers and government. To illustrate, in the Netherlands a foundational report was published by

the interdepartmental Behavioural Insights Network Netherlands (BIN NL), titled 'A Wealth of Behavioural Insights' (2018). This report was the first official BIN NL publication, and more importantly, the first centralized attempt to showcase the results of BI projects across the varied departments. All four pillars are reflected in the report; it starts out with a synthesis of insights from new behavioural economics, the showcased projects largely consist of RCTs testing the effects of nudge-interventions, and the overall message the report seems to make is one of basing policy more on (behavioural) evidence.

These four pillars can also be observed in the work done by the Australian Federal government. When BETA was first established they developed a Guide to Developing Behavioural Interventions for Randomized Controlled Trials (2016), in which they state that 'it's important to put real human behaviour at the centre of policy and programme design. Designing policy should be based on a sound understanding of human behaviour. This goes hand-in-hand with BETA's commitment to test those designs, building our understanding of what works and when we need to adapt our approach' (Ames 2016: 4). In order to better identify potential behaviourally-informed interventions the document also provides a high-level snapshot of some key behavioural science research including the use of reminders and personalization. It also encourages policy makers to draw on many of the nudges noted above such as default rules, social norms and pre-commitment strategies. The emphasis on behavioural science and nudges is also reflected in BETA's recent 'collection of stories about the work of [BI] units across government', in which it claims that 'behavioural insights offer ways to devise elegant and simple solutions to problems which can't be solved with traditional assumptions, wisdom or tools. The solutions are often modest and cost-effective, often yielding an impact far greater than more expensive policy options' (BETA 2018a). Of the 11 Federal governments interventions, 10 were evaluated with an RCT or framed field experiment, and 10 of the 11 could be classified as nudges (BETA 2018a). Lastly, also the focus on 'what works' is echoed in BETA's official documentation. The team for instance states on its website that it has a 'commitment to test… building our understanding of what works and when we need to adapt our approach.' (BETA n.d.-a)

It should be noted that this shared behavioural science-based government narrative, while consistent, remains a broad and abstract aspect of the field's identity and speaks more closely to the field's aspirations than its backstage practices. As we will see later, this seemingly fixed and coherent narrative can become operationalized in diverging ways. At the backstage, portrayed consistencies can turn into areas of contingency. Before we get to these contingencies however, we first want to point out aspects which still do seem consistently present in experts' daily work. As derived from our

ethnographic observations, specific sets of artefacts – in the form of heroes, role models and classic writings, and analytical and instrumental tools – appear to be consistently informing how experts operate. These artefacts are outlined below.

***Shared heroes, role models and classic writings***

One dimension of this shared set of artefacts related to the key writings and role models that have come to inform the field. An illustration of the importance of these core texts can be seen in BETA's establishment of its own 'library' for staff including copies of *Nudge* (Thaler and Sunstein 2008) and *Thinking, Fast and Slow* (Kahneman 2011). As noted by John (2018) and observed in our own research, leading academics have themselves become influential 'human artefacts' for the field. The Nobel Memorial Prizes of Kahneman and Thaler also are appropriated as key events in the field's history. These academic role models do not just exert their influence from a distance but have close ties to the field, giving advice, collaborating and presenting keynotes.

BIT UK has also acted as a role model to BI experts. With its pioneering work on using behavioural science in a policy context, it has generated models, frameworks, acronyms and even defining key terms that have come to dominate the field. Both in Australia and the Netherlands, it seems that many of BIT UK's projects and reports – particularly 'MINDSPACE' (Dolan et al. 2010), 'Test, Learn, and Adapt' (Haynes et al. 2012) and 'EAST' (Service et al. 2014) – have been of major influence on practices elsewhere. The use of the term 'behavioural insights' is a noteworthy sign of this influence, which is not only used by other native English-speaking countries but also many others, including Japan, Qatar and the Netherlands. In the case of BETA, it can be observed that some of its early initiatives were inspired by the work of BIT UK, for instance those recreating their energy labelling trials (BETA 2018b; BIT UK & Department of Energy and Climate Change 2014) and trials targeting the over-prescription of antibiotics (BETA & Behavioural Economics and Research Team 2018; Hallsworth et al. 2016). BETA has also followed BIT UK in exploring the issue of unconscious bias (BIT UK n.d.-a; Hiscox et al. 2017). In the Netherlands the influence of BIT UK can be observed by several members of the Behavioural Insights Network Netherlands (BIN NL) making a company visit to BIT UK, while the Dutch City of Amsterdam collaborated with BIT UK in organizing a BI course for employees in the communication department.

***Shared family of analytical and instrumental tools***

BI experts are further united by the use of specific tools and frameworks for analysing and designing policies. These include tools which outline how to develop a project, analyse behaviour and select instruments. Each iteration of these tools exhibits its

own variations and subtleties but the overlap between them seems strong enough to view them as a close family of tools.

One constitutive framework is a particular step-by-step way of setting up policy projects. While there are differing versions of these step-by-step plans, they all reflect a belief in the traditional rationalist policy cycle, moving from problem diagnosis to designing, testing and delivering solutions (Lindblom 1959). BETA, for instance, intends to set up each of its projects according to a four stage project development process. This process begins with a 'discovery' stage where research is undertaken to assist an understanding of the context and to gather information on the target population and their behaviours. This is followed by a 'diagnosis' stage where data and materials are reviewed to define the behavioural problem. This is then followed by the 'design' and then 'delivery' of an RCT (Ames & Hiscox 2016). In the Netherlands, we observe nearly equivalent versions. The BIT at the Ministry of Economic Affairs and Climate Policy (BIT EZK) follows these four phases: problem analysis, context analysis, possible interventions, testing. Similarly, the BIT at the Ministry of Infrastructure and Water Management (BIT IenW) refers to these stages: unravelling, designing, experimenting, monitoring and evaluating (BIN NL 2017b).

The use of nudges is also universal. Moving beyond the narrative, in BETA's case, all the projects that are currently publicly available on its website can be considered nudges (BETA n.d.-b). In the Netherlands, an update report by BIN NL (2018) showcases 14 projects which have been realized by the various departments, 10 of which include nudge-interventions, while the others tend to be centred on the stage of behavioural analysis or lack specific information about instruments. Admittedly, there are cases in which experts occupy themselves with the design of laws and financial incentives, but these are rare and even there the approach seems to be to complement these instruments with nudge-elements – for instance, trying to make a certain financial incentive more 'visible'.

These shared narratives and artefacts are important because they draw practices together for the purposes of future action. They serve to embody practice and allow it to be knowable by others. 'Indeed, it may even be the very existence of the object, its normal presence that leads actors to think and act on, with, through or around it: the artefact requires the practice, which in turn requires the artefact' (Freeman & Maybin 2011: 165).

## 6.4 Contingencies in the Behavioural Insights field

The analysis now shifts to areas of contingency in the field. These contingencies are where we can see how BI concepts are translated to suit different institutional contexts. Some of these contingencies speak to structural differences, e.g. to how experts are organized and which roles they take on. Although certainly important, these factors do not concern the core identity of the field and are not discussed here. However, it is with respect to methodology – and epistemology underneath – that we notice a deeper and more meaningful kind of translation, with practices that seem at odds with the field's frontstage narrative. These areas of contingency relate to how knowledge is actually dealt with in concrete situations, and as such are not yet visible in the field's frontstage performances. We identify three types of contingencies at this methodological level, which are detailed below.

***Behavioural Insights 'the method' vs. 'the theory'***

A first aspect of contingency concerns the difference between BI as *method* versus BI as *theory* (Einfeld 2018). BI experts appear to work along a spectrum which, at one end, highlights a focus on experimental methods, driven by a desire to find out 'what works', while the other is focused on applying theories of human decision-making drawn from a broad range of behavioural science disciplines. While plenty of practices seem to celebrate BI for the combination of theory and method, at the same time we can see particular teams clearly leaning towards one or the other. For example, during fieldwork with BETA, a strong focus on method over and above the use of BI theory was observed. Projects that lacked well-defined, measurable and quantifiable behaviours were considered outside BETA's remit. One of the staff even expressed their belief that 'the theory' was being used to help promote 'the method':

> I feel very strongly about that part of what we do. And then in some ways I think it's like a revolution, is the testing, is the randomized controlled trials, kind of using behavioural insights to get the evidence base stuff into policy. Not that that parts not important but they go hand-in-hand, I feel like we are piggybacking kind of off the popularity of behavioural insights, to be like 'also it's a super good idea to test your ideas'.

The case of the Dutch School Canteen Brigade, linked to the Netherland Nutrition Centre, tells a different story however. This unit aimed to improve the healthy eating behaviours of students in secondary schools. It has been relying on a host of theories, including the well-known experiments conducted by the alleged scientific 'food guru' Brian Wansink, revealing the power of framing, proximity and convenience in food choices (e.g. Wansink et al. 2006). Some of these theoretical findings, and the type of

interventions that were trialled to arrive at these findings, were straightforwardly applied without running RCTs to find out 'what works' locally. For instance, healthy types of bread and fruit would be placed more conveniently, and water would be made more available through the installation of nearby tap water points and the placement of jugs filled with tap water at the canteen counter (see Feitsma 2019). While this practice seems prototypical – dealing with a classic 'nudge site': the cafeteria (see Thaler & Sunstein 2008: 1-3) – it nonetheless represents a clear split from the earlier described practices like BETA. While theory-focused experts were more likely to adopt and imitate interventions drawn from previously successful trials, method-focused experts focused on replicating trials, acknowledging that local contextual differences can determine whether interventions are successful or not. This deepens the split between BI as method versus theory as it may make those who lean towards 'the method' increasingly sceptical of established behavioural theories, particularly given the current replication crisis within the fields of psychology and medicine. They may become 'theory-less', only taking local experiments as a valid source of knowledge about 'what works' – despite whatever track records of theorizing and evidence building exist.

In addition to differences in leaning more towards BI as method or as theory, we observed that BI experts differed more specifically in their use of methods and theories. These contingencies can be understood along the spectrum of (both methodological and theoretical) purism versus pluralism.

***Method Purist vs. Pluralist***

Our observations uncovered a second area of contingency with regards to the types of methods experts use and see as valid, which could broadly be described as *method purist versus pluralist*. On the one hand, there is a group of experts who hold on resolutely to the use of RCTs as the go-to method for evaluation. For them, RCTs possess an exclusive epistemic authority in making possible claims about 'what works'. BETA offers a clear example of a member of this group. Its Research Director was an economics professor from Harvard who was openly passionate about the value of running RCTs. For instance, in a presentation to the Institute for Public Administration Australia it was noted that:

> [the] head of the Commonwealth's relatively new behavioural economics unit thinks it is "sort of scandalous" that all public policy is not rigorously tested by default before it is rolled out. Experimental trials – randomised controlled trials wherever possible – are simply part and parcel of the behavioural interventions known as nudges that are fast becoming a standard tool for policymakers across the world. (Easton 2016)

This is not to say that there was not a degree of friction on this point in practice. For some BETA members, methodological satisficing was considered to be critical to being an employee of the public service. Some participants also expressed concerns that taking an 'RCT or the highway approach' might risk them being cut out of big policy decisions or projects as one interviewee stated: 'many departments really just want help solving their problems in a behavioural way not necessarily wanting a trial. By demanding a trial [we might be] missing opportunities.' Moreover, in practice BETA was often required to satisfice on their ideal trials. In their first publicly reported trial, in partnership with the Australian Public Service Commission, BETA undertook an RCT, 'the best way of telling if a policy is working'. However, they later noted that it could better be called a 'framed field experiment' and addressed the limitations of undertaking this study in a hypothetical environment as opposed to a real recruitment situation (Hiscox et al. 2017).

This observed satisficing of BETA, running 'framed field experiments' instead of pure RCTs, does however not nearly compare to the much greater degrees of methodological satisficing and pluralizing that can be observed in various other BI practices. By pluralizing we mean that experts replace the rationalist hierarchy of evidence that is tied to evidence-based policymaking (Cairney 2017), with RCTs as the gold standard, for a more horizontalized outlook in which multiple methods, each in their own ways, provide valuable information for policy design. An interesting example of such methodological pluralism is found in the work of an expert at the Royal Netherlands Army, who explained that he tries to find out 'what works' by simply putting ideas into practice and then observe and assess whether they work (see Feitsma 2019). Or one might look at an expert working in a secondary school, who claimed that partly due to time and resource constraints most of his behavioural designs were justified based on 'feeling' (see Feitsma 2018a). As a last example, we zoom in on an expert at Rijkswaterstaat, a Dutch executive agency. While recognizing the causal explanatory power of RCTs, he also distanced himself from a purist RCT advocacy:

> Well, that [running RCTs] is not really my thing, but it's necessary of course if you want to prove certain things… that it works… and to which extent it works. But it's difficult and costly to set up such an experiment in the, let's say, cruel-and-dirty outer world (…) and often it's about getting the precise effect and the quantification thereof. And then I think… there has to be a significant effect… and you can also observe that in different ways than through (…) very pricy and time-consuming trials. And that's often possible. [For instance] when [referring to an ongoing project about waste collection] those people during their collection activities notice

> something like: "Hey, this works, this really makes a big difference." Or: "We're collecting so and so much more." So you're going back to [relying on] estimations and key indicators… you can say something about those fairly quickly. I understand that in some cases you've really got to go to the RCT but that won't always be the case I think.

Besides downplaying the value of RCTs, being perceived as costly and technically difficult procedures providing disproportional degrees of rigor, the pluralistic outlook of this expert reflects in his openness to alternative evaluation methods, like estimations based on practical experience. He also noted the importance of bringing in the expertise of the local community and organizing local creative brainstorms together as a method to design policies. For instance, he used the 'Synectics' technique, a creative methodology that works with metaphor, imagery and the absurd in order to come up with novel ideas. Such alternative evaluative techniques are rarely mentioned in the prevailing BI discourse, which clearly prioritizes scientific over experiential knowledge with its core premise that '[i]ntuition is not evidence' (Wilson & Juarez 2015). The contingency observed here seems to run deep, and raises questions about whether the inherent tensions in the idea of a methodologically pluralist BI practice can be resolved. How pluralist can BI become before it stops being BI? Where is the methodological boundary?

***Theory Purist vs. Pluralist***

A third broad area of contingency concerns the theoretical bodies of knowledge that experts draw from, specifically linking to the split between *hardcore behavioural science versus other types of disciplinary knowledge.* As BI is rooted in behavioural science, all experts relied to some extent on behavioural science, particularly on behavioural economics and the behavioural and cognitive strands of psychology. However, some experts were more strictly tied to particular disciplines than others. For instance, the Team Behaviour Change of the Dutch Tax and Customs Administration, at the time employed by four social psychologists, appeared strongly wedded to social psychology. This commitment was reflected by the trialling of interventions that came from social psychology, such as Robert Cialdini's (1984) work on social influence. Many of its interventions focused on aspects such as wording, colour, distance, sequence and presentation, representing the prototypical behavioural science literature with its emphasis on the many 'seemingly irrelevant' factors within the immediate environment that impact individual behaviour (Thaler 2015).

The case of BIG'R, a centralized unit at the municipal level with close ties to Erasmus University Rotterdam, however tells a different story. BIG'R employed and collaborated not only with psychologists, behavioural economists and self-professed 'specialist[s] methodology and BIT-research' (BIG'R 2018) but also with ethicists,

sociologists, anthropologists and public administration scholars (Prij 2018). Its efforts in organizing public debates about the ethics of public behaviour change reflect a political philosophical orientation as well. Additionally, it is interesting to note how one of BIG'R's heads introduced the team's work methods and core assumptions. When talking about nudging and its corresponding concept of humanity, she did so from an explicitly sociological outlook, which is rather unusual in the field:

> Crucial is the insight that environmental factors are much more significant in shaping individual behaviour than often is assumed. (…) Too much reasoning starts from the idea of self-management and self-control. That is deceptive, because the environment in which you grow up to an important extent determines your choices and your behaviour and your opportunities to develop. This is something I have to explain well, and I am referencing here to the well-known sociologist Bourdieu, who claims that the interaction between society and individual occurs on different terrains, and that people need economic, social and cultural capital in order to shape that interaction (Prij 2018: 47).

The quote above shows a recognition that behaviour, especially in the case of wicked policy issues, cannot simply be changed by redesigning the immediate environment at the micro-level. Sustainable change instead has to come from deeper interventions into the fabric of society, targeted at real changes in the economic, social and cultural capital of policy subjects.

The deep translation we have observed raises puzzling questions about the field's boundaries. How far can a BI practice go in drawing from either seemingly conflicting disciplines and remain BI? How should we understand such alternative, heavily adapted practices – have they simply given way to the driving forces of local context, or do they form alternative but nonetheless compatible versions of BI, moving the field in new directions? Where do the boundaries of BI actually lie? These 'boundary issues' will be explored further below.

## 6.5 Conclusion and discussion

### *Mapping the boundaries of Behavioural Insights*

In this chapter we have analysed the increasingly popular BI movement, exploring how its agenda has been moving across the global policy sphere. We have spoken about several key ways BI has manifested in two different contexts. What we have observed is that experts within this field, despite all operating under one umbrella concept, are at the same time translating their practice to suit local traditions and needs. While they share a similar narrative, and draw on similar role models,

analytical frameworks and instruments, they diverge significantly on how they handle knowledge in practice. Table 6.1 summarizes the main aspects of consistency and contingency that we have observed in the field.

**Table 6.1: Consistencies and contingencies in Behavioural Insights practice**

| | *Patterns* | *Examples* |
|---|---|---|
| *Consistencies* | I Shared frontstage narrative<br>II Shared family of heroes, role models and classic writings<br>III Shared family of analytical frameworks and instrumental tools | I The widespread dominance of 'what works' lexicon within official discourse, and the repeated call for more RCTs in policy analysis, for instance see BIT UK's 'Test, Learn, Adapt' paper.<br>II The works of Thaler, Sunstein and Kahneman. The global influence of BIT UK.<br>III The universal adherence to the rationalist 'diagnosis-design-delivery' policy cycle. The widespread adoption of MINDSPACE, EAST and the nudge-toolkit. |
| *Contingencies* | I Behavioural Insights 'the method' vs. Behavioural Insights 'the theory'<br>II Method purist vs pluralist<br>III Theory purist vs pluralist | I BETA, committed to RCT-based projects vs. the School Canteen Brigade, straightforwardly applying well-known social psychological theories without running local RCTs.<br>II BETA, largely evaluating projects through RCTs vs. solo expert at Rijkswaterstaat, deciding about instruments on the basis of a broad palette, including RCTs, quick analyses, practical experience, creative thinking and brainstorms with local stakeholders.<br>III Team Behaviour Change, consisting solely of trained social psychologists vs. BIG'R, also employing and collaborating with anthropologists, ethicists and political scientists. |

***Reflections on Behavioural Insights from a policy translation perspective***

Our observations about the boundaries of BI raise several interesting discussion points regarding policy translation. To begin with, the individual experts have demonstrated an investment in expanding BI in an attempt to help governments develop and implement better, evidence-based policy. Much of this investment so far has hinged on the concept of diffusion – BI as a concrete idea that can be shared across governments, between countries and within departments including through the communication of mnemonics and checklists (John 2018: 77). However, our observations challenge such a realist-objectivist diffusion way-of-thinking and suggest that BI 'moves' in a way that is more closely coincides with the process of translation. Originally proposed by Latour (1986), translation is more interpretive and

continuous, focusing on how individual actors interact with the concept and change it 'into something completely different as they [seek] to achieve their own goals' (Latour 1986: 268).

In fact, our observations of the field's developments may even reflect a more complex idea of translation. We saw BI practice as composed of several individual components unified largely by an overarching narrative but none of which appear to be reliant on the other. This combination of distinct, and plausibly independent, technologies clearly demonstrates the use of artefacts – shared narratives, frameworks and instruments – as boundary objects to facilitate translation across time and space. As discussed by Star and Griesemer (1989: 393):

> boundary objects are objects which are both plastic enough to adapt to local needs and the constraints of the several parties employing them, yet robust enough to maintain a common identity across sites. They are weakly structured in common use, and become strongly structured in individual-site use. These objects may be abstract or concrete. They have different meanings in different social worlds but their structure is common enough to more than one world to make them recognizable, a means of translation. The creation and management of boundary objects is a key process in developing and maintaining coherence across intersecting social worlds.

Building further on this conception we can turn to the work of Joan Fujimura (1992) who notes that these boundary objects can in turn be effectively bundled together to create 'standardized packages'. These packages bring together theory and methodology and construct a stable definition. This facilitates coordination, collaboration and encourages the production of coherent texts. This is essentially what she refers to as a 'grey box', wherein individual components are 'visible' but shielded from critique by the broader narrative which combines them. We would argue that BI represents an example of a standardized package. The abstract theory uses a myriad of behavioural science concepts, largely captured under the hard-to-challenge idea that policy should take into account human behaviour, combined with the specific, standardized technologies of A/B tests, RCTs and nudge.

Following on from this, we can see how the spread of BI as a standardized package has to some extent immunized its individual components (i.e. nudges, RCTs, behavioural insights) from challenge. A nudge can be disputed on ethical claims, but if it is being tested then any issues should be revealed and, in turn, stopped. And in absence of a trial, how can one know if the policy is working? This brings us back to the BETA team member who stated that 'they go hand-in-hand, I feel like we are

piggybacking kind of off the popularity of behavioural insights, to be like "also it's a super good idea to test your ideas"'. These practices need not coexist: we have seen examples of nudges being undertaken without RCTs and RCTs are able to be done without nudges, but for now their combination within the field of BI acts to strengthen the component parts beyond their capacity individually.

This is particularly important because the selection of policy instruments is not a passive or politically neutral decision. The use of nudges for example, as noted above, has raised a number of moral and democratic concerns. However, because much of the work in selecting an instrument tends to be disguised as technical or scientific, or represented by the rational or instrumental application of an instrument to a policy problem, they are largely depoliticized (Simons & Voss 2018; Lascoumes & Le Galès 2007). By bringing all of the ideas together into a package these experts have been able to limit their exposure to the overtly political or problematic aspects of nudging in addition to pushing the boundaries of what a BIT offers.

Connecting this back to the translation literature, this tells us something about the power of narrative in facilitating the translation of global policy agendas. The successful adoption of BI practices depends on the ongoing 'connective' work that is being performed at the frontstage of the field, gluing individual components together into a universal, 'movable' package. Interestingly, and this is part of the complexity of how BI gets translated, this degree of connectedness is not so much observable at the backstage. Our ethnographic fieldwork reveals a landscape which, although imagined more coherently, ultimately is rather heterogeneous. Deep splits in the field exist with respect to actual knowledge use, and reflect paradigmatic differences. In terms of theory use, most experts subscribed to the individualist behaviour change paradigm that also prevails at the frontstage – but others drew from a far wider disciplinary pool. In terms of method use, we observed experts for whom nudging in practice was largely about policy imitation, while on the other side RCT-advocates saw the lack of testing as 'scandalous'.

***Valuating Behavioural Insights' translation: boundary issues or ambiguity tactics?***
Having demonstrated how BI is translated at these two sites, particularly with respect to its seemingly strict frontstage yet underneath loose backstage boundaries, it is interesting to reflect on why these translation dynamics may have occurred and what function they might play. How can it be that such different practices can be operating under what is still perceived as one coherent hood?

We can offer some initial thoughts on this. One thing that has to be considered is the vast scope of this trend. As BI is traveling across nations and continents, it is

unsurprising that it is to some extent doing so in a fluid way. The time horizon may play a role as well. The field is at a relatively early stage of development and institutionalization, leaving its boundaries still in the process of being defined. Moreover, behavioural science is a broad field and perhaps this broadness in turn widens the field's boundaries and makes fragmentation, in which subgroups form drawing on particular insights or particular technologies, more likely.

It ought to be recognized that experts may reap important benefits from having more porous backstage boundaries while at the same time possessing well-defined boundaries at the frontstage. Such a hybrid constellation allows space for the development of pluralist and alternative approaches without evoking conflict and without appearing boundaryless and incoherent at the frontstage. It creates ambiguity, a form of hypocrisy (Brunsson 2007) in particular, which has generally been considered a strength when working with policy and government. It enables support from broad and ideologically diverse quarters (Stone 2012). This type of ambiguity also allows for negotiation and compromise, as 'legislators can satisfy demands to "do something" about a problem by passing a vague statute with ambiguous meaning, then letting administrative agencies hash out the more conflictual details behind the scenes' (Stone 2012: 158). Bringing this back to how BI is 'performed', we observe a frontstage that draws legitimacy from telling a rigid and clear-cut story about how policy *should* be designed in ideal circumstances, while simultaneously strategically leaving unaddressed how that story is to be realized in the often obdurate and clashing world of policy practice.

There may however also be a risk attached to the field's broad narrative and backstage fluidity. The tactical use of ambiguity may also backfire into a disorienting situation in which the field suffers from 'boundary issues'. This risk has not gone unnoticed. A recent paper by Lepenies et al. (2018: 177, italics added) noted that

> [e]xplicitly demarcating and acknowledging the limits of the behavioural sciences – both empirically and normatively – *may reduce the scope of the application of 'behavioural' tools to policy-making, as well as the political influence of such approaches*, possibly beyond a point that the authors and other proponents would find desirable. Yet it might be a long-term strategy that is in their interest. Claiming more for (behavioural) science than is due puts the credibility and legitimacy of behavioural approaches at risk.

There were times when the behaviour experts in both of our studies also reflected on this, particularly when considering their interactions with policymakers. On several occasions during the fieldwork, one author was informed that while there was noticeable enthusiasm for BI within government, it was unclear whether this was

because of any specific trait or because it was perceived as the most effective way to get a policy issue on the agenda. One senior member of BETA lamented the fact that several of the projects they were being asked to undertake reflected people's desire for solutions to general problems rather than any particular interest in the behavioural agenda. She said at the time, 'most of the time it was "can you solve my problem", not "here is a behavioural economics problem"'. Even more strikingly, in an interview with a long-serving senior public servant, it was shared that

> [a] new policy innovation as a bit like a five-year-old's soccer game. So there's the ball and there's everybody crowded around the ball, where the ball goes everybody runs, so that's what behavioural insights is like in policy terms now. If we don't embed it properly into the policy cycle and policy thinking, it will be like that until everyone decides it's not working because all you get is a ball with a whole bunch of people crowded around it, which is really boring to watch, and they will move onto the next thing, and I've seen that happen in a few things now (..) so (..) we need to help people work out when it's a good thing to use and when it's not.

Thus, to conclude, stricter boundaries may have their value as well. At this early stage of time, it is only logical that BI is still 'in the making'. However, the question remains whether it will eventually move beyond that phase and flourish into a more uniform practice that walks it talk more fully. In any case, policy scholars have a vital role to play with regard to these questions, firstly by keeping track of the field's front- *and* backstage developments. Secondly, realizing that the field's current contingency implies the possibility of change, scholars can help transform the field, pinpointing blind spots and suggesting new avenues forward.

*Chapter 7*

# Institutionalizing the Behavioural State

*Public policy design takes place in a complex 'policy swamp' that is not easily analysed, let alone controlled. Nonetheless, recent scientific advances in understanding human behaviour have led some to believe there is a way out of this swamp. A 'Behavioural Insights' movement has emerged, pushing a seemingly neorationalist strategy that clashes with the hitherto incrementalist strategy of policymaking. This chapter investigates how upcoming behaviour experts in Dutch government grapple with this clash, based on long-term ethnographic fieldwork. The chapter points out that these behaviour experts, despite their clear-cut rationalist impression, in the backstage take on the challenge of negotiating competing institutional logics.*

---

Chapter 7 is published as an original research article by Joram Feitsma in Critical Policy Studies.

Feitsma, J.N.P. (2018). 'Rationalized incrementalism'. How behavior experts in government negotiate institutional logics. *Critical Policy Studies,* advance online publication. doi: 10.1080/19460171.2018.1557067

## 7.1 Introduction

Worldwide, governments are making a 'behavioural turn' (e.g. Strassheim et al. 2015; Leggett 2014; Whitehead et al. 2017; John 2018; Malecka & Lepenies 2018; Schubert 2017). Increasingly, findings from the new school of behavioural economics in particular – showing how people systemically deviate from *homo economicus* behaviour – are being considered to achieve policy goals, whether that concerns sustainable environments, healthy lives, or compliance with security laws or tax regulation. This behavioural turn has materialized in the rise of a 'Behavioural Insights' movement (Dolan et al. 2010; John 2018). Apart from writing influential reports like Mindspace (Dolan et al. 2010), this movement develops savvy 'nudges' (Thaler & Sunstein 2008) that go with the grain of human behaviour, while drawing on rigorous Randomized Controlled Trial (RCT) methodology to test the causal effects of their interventions. Behavioural Insights' core aim thus is to align public policymaking with the psychology of human decision-making, making effective policies based on experimental evidence.

Despite the global achievements of forerunning behavioural units (John 2018), a successful deep institutional change like the behavioural turn is far from guaranteed. Governments are at this moment still 'experimenting' with behavioural units and associated practice in relatively marginal capacities. Also, the establishment of Behavioural Insights is not a neutral project. It has evoked skeptics and critics across the policy-politics nexus who explicitly resist the behavioural turn. More implicitly, Behavioural Insights faces a systemic resistance to change within the public domain (Ansell et al. 2015), while eliciting a paradigmatic struggle between contrasting knowledge disciplines that is not necessarily in favor of the as of yet fledgling behavioural science perspective (although see Malecka & Lepenies (2018) on the 'scientific imperialism' of the behavioural sciences). An additional puzzle is that Behavioural Insights – despite its self-invented rhetoric of 'Radical Incrementalism' (Halpern 2015) – appears to hold a very *rationalist* outlook on policy design, attributing a central importance to scientific design and analysis in its quest to manage social problems (Lodge & Wegrich 2016). Decades of policy studies have shown that such an outlook is far from obvious. The policy world is actually rather *incrementalist*, and attributes a much more limited and messier role for science (e.g. Lindblom 1959). This evokes the paradoxical observation that while Behavioural Insights accentuates the bounded rationality of policy subjects, it still tends to assume unbounded rationality in policy designers and governmental processes.

In light of the abovementioned 'rationality paradox' (Lodge & Wegrich 2016), this chapter seeks to understand what happens when behavioural units try to inject their neorationalist ideas and methods into government organizations and policy

subsystems that seem to run counter to such a rationalist ambition. It explores how behaviour experts grapple with these conflicting (i.e. rationalist versus incrementalist) logics of policy design. Grounded in ethnographic fieldwork on Dutch governmental behaviour experts over the course of 4 years, I observed that these behaviour experts carry out a strong and unequivocal rationalist message in the 'frontstage'. In the 'backstage' however they attempt to harmonize their frontstage image with the 'real world' of incrementalist policymaking. The distinctive contribution of this chapter therefore lies in challenging the wholly post-political, neorationalist imago of Behavioural Insights from an empirically informed perspective. Drawing on observations from the field, it will be demonstrated how in the backstage a more hybrid, 'rationalized incrementalist' logic is operative.

## 7.2 The rationalism versus incrementalism debate

In the existing behavioural policy literature, there is little mentioning of the rationalism that appears to underpin Behavioural Insights at a deep level. Its rationalist outlook on policy mostly seems assumed and therefore remains unaddressed. This chapter however takes this implicit rationalism as its explicit object of analysis. Doing so, it falls within a small group of critical studies that have been scrutinizing developments of modernization, scientization, and rationalization in the field of behavioural policy (e.g. Lodge & Wegrich 2016; Strassheim et al. 2015; Feitsma 2019). It adds to this critical literature by approaching this rationality theme through the lens of the classic rationalism versus incrementalism debate within public administration science. A recap of this debate is presented below.

Policy design has generally been construed as an inherently messy phenomenon. This messiness reflects in well-known metaphors relating policy design to 'primeval soups' (Kingdon 1984), 'garbage cans' (Cohen et al. 1972), and 'swampy lowlands' (Schön 1983). At the same time, public administration has a long history of rationalist waves attempting to dodge the messiness of the 'policy swamp' (Parsons 2002). Under a rationalist view, policy design can and should be approached scientifically and synoptically, through analysing problems in all of their aspects, mapping out all possible solutions, and selecting and rolling out the most optimal solutions based on their identified consequences (Lindblom 1959). An early and prominent rationalist wave was the Anglo-Saxon positivist policy analysis movement in government that arose after World War II (Fischer et al. 2015). Since the 1970s, there was a new wave arguing for more 'research utilization' in policy (Weiss 1979; Newman 2016). Since the 1980s, the idea of 'social marketing' gained widespread popularity (Pykett et al. 2014), and since the 1990s another wave emerged in the form of the 'evidence-based policy' movement (Strategic Policy Making Team Cabinet Office 1999; Strassheim 2017). A most recent wave is the rising Behavioural Insights movement, whose

rationalist bend is highlighted by its evidence-based approach and radical advocacy of RCT methodology (e.g. Lourenço et al. 2016; OECD 2017; Haynes et al. 2012).

Against the rationalist logic, an incrementalist logic has been posited, memorably depicted by Charles Lindblom (1959) as 'Muddling Through'. This has sparked a longstanding academic debate about the rationalist versus incrementalist nature of policy design (Lindblom 1959; Dror 1964; Etzioni 1967; Smith & May 1980). In this debate, the rationalist view has been critiqued on several of its implicit and explicit assumptions, including its beliefs in: the unboundedly rational nature of the policy process; the possibility of synoptic analysis and radical policy change; the fixed nature of knowledge; the hegemony of causal knowledge; the solely instrumental role of evidence; and the appropriateness of a command-and-control policy delivery model (Lindblom 1959; Parsons 2002; Cairney 2017; Boswell 2017; Strassheim 2017). Table 7.1 (drawing from Lindblom 1959; Smith & May 1980; Parsons 2002) summarizes the main issues in which these logics clash.

**Table 7.1: Rationalism versus incrementalism**

| *Rationalist logic* | *Incrementalist logic* |
|---|---|
| - Policy design as a rational-analytical and value-free enterprise | - Policy design as a rationally bounded and political affair |
| - Comprehensive analysis of values, alternatives and outcomes | - Partial, economized analysis of values, alternatives and outcomes at the most |
| - Radically new designs | - Only incremental changes |
| - Outcome driven | - Process driven |
| - Emphasizes fixed, decontextualized, causal 'what-works' knowledge (Episteme) | - Emphasizes ambiguous, local, experiential 'how-to' knowledge (Phronesis and Techne) |
| - Assumes an uninterrupted way from science to policy | - Assumes a diffused and 'brokered' way from science to policy |
| - Requires a top-down, managerial, powering model for policy delivery | - Requires a bottom-up, fragmented, local puzzling model for policy delivery |

Importantly, Behavioural Insights cannot be wholly generalized as 'just another' rationalist policy wave. One aspect in which it seems to have embraced incrementalist critiques on rationalist views is its departure from a radical design theory. Instead, Behavioural Insights seeks to generate social change by working incrementally, underpinned by its promoted philosophy of 'Radical Incrementalism': 'the idea that dramatic improvements can be achieved, and are more likely to be achieved, by systematically testing small variations in everything we do, rather than through dramatic leaps into the dark' (Halpern 2015: 291). While with regards to this aspect of

making small steps (versus radical systemic change) Behavioural Insights indeed makes a shift toward incrementalism, it would be inaccurate to depict 'Radical Incrementalism' as incrementalist in the original Lindblomian sense. Even when making only one small policy change at the time, this strategy still adopts a strongly rationalist, scientized, engineerist policy approach toward the 'management' of social problems. This stands at odds with the incrementalist observation of a chaotic and politicized process in which little steps are made not through deliberate synoptic analysis but through highly satisficed decision-making resulting from negotiated, contingent, unplanned prioritization of certain values and interests over others (Lindblom 1959).

We might then view 'Radical Incrementalism' not so much as an accurate description of the actual institutional logic guiding behavioural policy, but more as a strategic politico-epistemological frame that helps the Behavioural Insights field to guise some of its rationalist traits. 'Radical Incrementalism' might even act as a 'magic concept' (Pollitt & Hupe 2011) to transcend the 'old' rationalism–incrementalism impasse. This chapter will however critically scrutinize this rhetorical frame by resituating it in the longstanding rationalism–incrementalism debate it has either explicitly or implicitly ignored. Through the lens of this classic public administration debate, we come to see the more balanced, mixed, 'rationalized incrementalism' that actually underpins Behavioural Insights.

## 7.3 Methods

This chapter is grounded in long-term ethnographic fieldwork on Dutch behaviour experts. I have studied these experts in their professional environments, for longer periods of time, trying to grasp their ways of making sense of the world, making observations, taking notes, and constructing a personal narrative from it (Rhodes et al. 2007). As of yet there is only a limited body of qualitative studies on how behavioural policy is developing (e.g. John 2014, 2018; Ball, et al. 2017; Jones et al. 2013; Jupp et al. 2016; Feitsma 2018a, 2019; Whitehead et al. 2017). More of such studies would seem needed to gain a deeper understanding of actual behavioural policy practice, especially as the literature currently tends to be dictated by either (overly) optimistic 'behaviouralists' reiterating the same type of generic ideas and success stories, or (overly) critical 'ethicists' offering rather abstract critiques (Whitehead et al. 2017). Qualitative, ethnographic approaches can help to go beyond the current 'trench warfare' type of debate, and build more empirically grounded accounts of behavioural public policy. These approaches can help to go beyond the fabricated 'frontstage' of behavioural institutions and retrieve insights about what happens 'backstage' in the field (Goffman 1959; Van Hulst 2008).

This chapter gathers its data from an overarching study on the emerging behavioural state, divided over four distinct research phases. The general research context is Dutch government, where various behavioural practices have been developed since 2009 in a relatively decentralized manner (for more detail on the rise of Behavioural Insights in the Netherlands, see, e.g. BIN NL 2018; Feitsma 2019; Feitsma & Schillemans 2019). From 2014 to 2016, I performed semi-structured interviews (with 35 experts across 20 organizations) and short-term observations (about 55 hours) of a wide range of Dutch behaviour experts at the level of central government. During this time, I was also involved in a local nudging project as an academic adviser over the course of ten months. In 2016, I worked as employee-ethnographer for a small Dutch behavioural team for a period of four months (from September to December 2016, totaling about 400 hours of participant-observation). From 2016 to 2017, I conducted another round of semi-structured interviews (with 15 experts across 11 organizations) and short-term observations (about 19 hours), this time focusing on behaviour experts in local government. In a last phase in 2018, I conducted a focus group with a mixed group of behaviour experts in order to validate concluding insights gained from the fieldwork. The findings presented in this chapter have also been discussed in-depth during this focus group.

I have thus far published about the emerging behavioural state from different theoretical perspectives, *inter alia* focusing on *everyday practices* (Feitsma 2019), *professionalism* (Feitsma & Schillemans 2019) and *modes of expertise* (Feitsma 2018a). This chapter adopts its own distinctive lens, looking at the particular *institutional logics* that underpin behavioural policy against the background of the long-standing rationalism versus incrementalism debate within public administration.

Regarding my sampling strategy: I specifically looked for self-proclaimed governmental behaviour experts who were explicitly and structurally making use of behavioural scientific knowledge. No further search criteria were adopted – for instance regarding professional background or organizational design – as the professionalism of behaviour experts was still in the process of being developed and already relatively fragmented – despite some shared core traits (Feitsma & Schillemans 2019).

Throughout the overarching study, interviewed and observed experts were generally made aware of my research aims, approach, and scope of study, typically during introductions and the process of negotiating access. At the same time, in practice the line between overtness and covertness in ethnographic fieldwork can become blurry, as awareness of the researcher's identity can fade during the fieldwork and time and space to make known the research agenda may be unavailable (Lugosi 2006).

An important question is whether my observations of Dutch behaviour experts are reflective of the wider international landscape. While Dutch experts are clearly informed by international role models (Feitsma & Schillemans 2019), numerous aspects of this context – for instance its traditional modes of expertise, public innovation culture, and particular history with 'evidence-based policy – are likely to shape behavioural policy practice in unique ways. Exploring the role of such contextual aspects in shaping varieties of behavioural policy practice forms an important avenue for future research (for an example, see Strassheim et al. 2015). A related question is whether I have really been able to uncover the better hidden backstage of behavioural practice. During interviews and short-term participant observations, I was a relative outsider which made it more difficult to be let into backstage processes. However, in some longer term involvements – collaborating in a local nudging project and working directly for a ministerial behavioural unit – I was part of daily backstage operations. At the same time, it has been my experience that the boundaries between the frontstage and backstage can be blurry. My use of the frontstage/backstage dichotomy should therefore be taken as a loose heuristic. One last question is whether the fact that I have been conducting research on behavioural public policy for the last four years, and have myself been embedded in a behavioural team, has led me to overestimate the significance of the 'behavioural turn'. It should therefore be kept in mind that in the wider institutional picture behavioural policy still remains marginal – despite the attention that the topic has received from numerous governments and the launch of new policy units (Whitehead et al. 2017).

## 7.4 The frontstage and backstage of Behavioural Insights

In his dramaturgical theory, the sociologist Goffman (1956) observes how people display different behaviours across different stages. An important distinction is that between the formal frontstage behaviours that serve as a form of 'impression management' (Goffman 1959) vis-à-vis the outside world, and the informal backstage behaviours that occur out of the public eye and allow for preparing for, reflecting on, and possibly deviating from the frontstage impressions. This frontstage/backstage distinction helps to make sense of my observations of Dutch behaviour experts. These experts tend to tell a highly rationalist story in the frontstage. In the backstage, however, a more complex reality becomes apparent in which they try to harmonize rationalism with incrementalism through various 'balancing acts'. This section will first describe the rationalist frontstage, after which the more hybrid backstage realities will be analysed.

***The rationalist frontstage***

In frontstage settings (e.g. public events, lectures, newsletters, strategic documents, or introductory conversations), behaviour experts express a strong rationalist ambition in which emphasis is placed on managing the behaviours of policy subjects through comprehensive analysis, rigorous *ex ante* evaluation, and systematic design. A first aspect of this rationalism is the experts' adherence to the traditional rationalist policy cycle, moving from problem definition to analysis to solution design to implementation to evaluation (Cairney 2017). These various steps in the rationalist cycle reflect in the type of step-by-step approach that the experts typically promote. For instance, BIT IenW, the behavioural team of the Ministry of the Infrastructure and Water Management, follows a comparable approach ('DOE-MEE') which entails four phases: unraveling, designing, experimenting, monitoring, and evaluating (BIN NL 2017). Such step-by-step approaches are generally presented as core aspects of behavioural practice, both in lectures and public reports (e.g. BIN NL 2017), but also during interviews. Table 7.2 (first theme, also see Appendix I) provides evidence of the behaviour experts' adherence to the rational policy cycle, seeking to shape policymaking into a systematic problem-solving exercise.

**Table 7.2: Evidence of rationalism**

| *Theme* | *Example* |
|---|---|
| Rationalist policy cycle | An interviewee at the Dutch Healthcare Authority, and who is part of a behavioural network, talks about her daily practice: 'That's actually pretty broad. Mostly… we're busy with looking how we can steer behaviour. That tends to begin with the development of a regulation strategy. Then we try to analyse certain structures in the behaviour, analyse the behaviour, and on the basis of that we also explore which interventions might be most effective to change the behaviour'. |
| Belief in readability and craftability of human conduct | Various experts talk about their behaviour change approach in terms of finding out 'what makes people tick' and 'pushing the right buttons'. They adopt different models to find out what these buttons are. One interviewee at the Royal Netherlands Army distinguishes between three 'buttons', being greed, fear, and sociality. In his view, these 'buttons' can be operationalized into behaviour change interventions, which he called 'weapons of influence'. Another interviewee, working at Rijkswaterstaat, distinguishes between three other 'buttons': opportunity, motivation, and capability. He notes how all human behaviour can be explained with the psychological theory of Maslow's pyramid of human needs, from basic needs to social status-related needs to the need for self-actualization. |

| | |
|---|---|
| Hegemony of RCTs | A key report in the field describes one of the core lessons learned when applying behavioural insights. Running field experiments is seen as a core pillar of behavioural practice: 'Testing measures in practice is crucial. Something can look good on paper, but the question is always whether it actually works in the relevant situation. (…) Use field experiments: test the interventions on a small scale in order to see what works and what doesn't. It is thereby important to measure what exactly the behavioural effects of the measure are, and what factors influence those behavioural effects. The rapid developments in data analysis and big data facilitate this.' (BIN NL 2018: 44) |

Beyond this frontstage promotion of step-by-step problem-solving methods, another aspect of rationalism in the world of behaviour experts relates to their both implicit and explicit beliefs about the 'readability' and 'craftability' of human behaviours. While one might expect these experts to emphasize a degree of analytical uncertainty in their work, moving away from clear-cut *homo economicus* assumptions while incorporating the complex and arcane territory of the human unconscious, expressions of such uncertainty are rather scarce at the frontstage. Rather conversely, it seems as recent behavioural scientific discoveries have strengthened the experts' analytical confidence. This for one reflects in the type of behavioural analytical models they employ. These models tend to categorize the origins of behaviour into several dimensions, for instance capability, opportunity, and motivation. Such models implicitly assume the possibility of clean-cut analysis in which the origins of behaviours can be determined with certainty and can be neatly fit into a given analytical framework. The idea furthermore is that when is the specific drivers and barriers behind behaviours are determined, those behaviours can be altered by making targeted interventions that address those drivers and barriers. From their increased confidence in their ability to analyse, or 'read', citizen behaviours, behaviour experts thus also gain more confidence about their ability to 'craft' citizen behaviours. The perceived simplicity of reading and crafting behaviours is symbolized by a prevailing mechanistic language, framing behaviour change as finding out 'what makes people tick' and 'pushing the right buttons'. Table 7.2 (second theme) gives examples of the expressed belief in the readability and craftability of human conduct.

A third aspect of rationalism is found in the experts' common adherence to a strict hierarchy of evidence, with RCT-based knowledge as the 'gold standard' of evidence. This hegemony of RCTs is one of the quintessential elements in the 'Radical Incrementalism' philosophy, wanting to improve public services 'one RCT at the time' (Halpern 2015). Generally, behaviour experts will show their support for

experimental evaluation. It is telling that the first update report of the Behavioural Insights Network Netherlands (BIN NL 2018), one of the more institutionalized Dutch networks of behavioural expertise, showcases its recent activities with the results of a series of field experiments. The emergence of behavioural policy thus aligns well with the RCT-oriented evidence-based policy trend which has been in vogue for some longer time already. Table 7.2 (third theme) shows another observed example of the attributed authority to RCTs.

***Balancing acts backstage***

In contrast with the strong rationalism propagated at the behavioural policy frontstage, in backstage-settings I observed behaviour experts seeking ways to alleviate the tensions that follow from carrying out a rationalist message in a not-so-rational policy environment. These ways, some occurring deeper in the backstage than others, showed behaviour experts distancing themselves from several rationalist assumptions such as: an 'unbrokered' and apolitical science-policy nexus; the need for and possibility of comprehensive analysis; a predominant focus on outcomes; and the hegemony of 'what works' knowledge (see Table 7.1; Lindblom 1959; Cairney 2017; Newman 2016; Boswell 2017; Parsons 2002). This is not to say that behaviour experts fully lost their rationalism backstage. Rather, they tried to live up to their rationalist ideal of basing policy design on scientific evidence (Newman 2016) as much as possible, while at the same time acknowledging existing incrementalist forces that run counter to such an ambition.

This negotiation between institutional logics reflects in particular 'balancing acts'. In the following sections, I will highlight four particular backstage balancing acts that emerged as salient in a process of linking my field experiences with the sensitizing concepts from Table 7.1: (1) knowledge brokering; (2) focusing on outputs; (3) analytical satisficing; and (4) horizontalizing the hierarchy of evidence. The first balancing act of knowledge brokering refers to the ongoing negotiation and merging between competing logics that happens at the *procedural* level. It refers to the blurred and politicized relationships of behaviour experts with their outside worlds, which cannot be successfully managed by following a rationalist logic alone. The other three balancing acts are tied to the negotiations between logics that take place at the *substantive* level; they are about how the work methods, theories, and criteria of behaviour experts are adapted in light of competing logics. In an overarching sense, all four balancing acts follow the same mechanism: strong rationalism is dismissed and a degree of incrementalism acknowledged, while rationalism is preserved as much as possible.

*Knowledge brokering*

The specific mechanism behind the first balancing act, knowledge brokering, goes as follows. A rationalist logic assumes an automatic flow from evidence to policy (Cairney 2017). Behaviour experts, however, both implicitly and explicitly reject this assumption. They recognize that evidence needs to be properly introduced, moved around, translated, and (re)contextualized before it gets 'used'. Therefore, they adopt the role of a 'knowledge broker': a mediator between behavioural science academia and the policymaking world (Knight & Lyall 2013; Hoppe 2010; Ward et al. 2009). This knowledge brokering role becomes apparent in at least three aspects. A first aspect concerns the experts' efforts in forming networks of expertise and circulating knowledge throughout the policy process. A second aspect is their support given to others regarding the concrete application of behavioural knowledge A third aspect is their strategic work in seeking to build political-administrative legitimacy. Table 7.3 gives more evidence of these three ways of knowledge brokering (for a more detailed analysis of behaviour experts as knowledge brokers, see Feitsma 2019; for more empirical evidence see Appendix I).

**Table 7.3: Evidence of knowledge brokering**

| *Theme* | *Example* |
|---|---|
| Smoothing the circulation of knowledge | One interviewee notes that her team operates as a loose network, consisting of policy advisers, knowledge institutes, behavioural scientists and expert practitioners. Her daily work consists for a large part of connecting and interacting with people. Her perceived added value is largely due to her built-up combination of 'know-what' and 'know-who': I have many interactions, a large part of my working hours is filled with interactions. Yes, you could say my job is interactions. And making use of the network to supply policy makers with content. But I also dive in the subject and in the studies myself. I can only think along with policy if I do my homework myself. The advantage is: I know both the content and the people. I also know what's necessary and what it is the people do and can provide.' |

| | |
|---|---|
| Facilitating concrete application | A ministerial behaviour expert explains the importance of facilitating others in the concrete application of behavioural knowledge. His team has developed its own 'DOE-MEE tool' to do so. 'Policymakers certainly realize that behaviour is an essential factor in achieving successes. But they've got their own professionality, which is often judicial, technical, or economical. That's the basis from which they develop ideas. That's what they hold on to, because people know how to do that. That's what they studied for. That's what they've gained experience in. That our behaviour also plays a role, and that it also works unconsciously, they realize all of that. Only the how-question hasn't been answered. Of: how do I then deal with it? How do I ensure that it becomes part of my policy process? This is where we step in with our DOE-MEE approach. We try to answer the how-question. That's why we also called it a "tool".' |
| Politico-strategic work | A municipal behaviour expert elaborates on a key challenge, which is not just a substantive challenge but also a political challenge of having to authorize oneself within the established policy system. 'The burden of proof is always important when telling your story. Is it scientifically proven to actually work? That's an important question and always important for political figures like my alderman. Because yea, (…) he has to make a case for why we're doing what we're doing. The scientific substantiation can be difficult. Although with telling examples and qualitative substantiation you can often come a long way.' |

These three ways of knowledge brokering reflect a shift toward incrementalism. The knowledge circulating function appears more in line with an incrementalist logic, acknowledging the distorted path from science to policy. Similarly, the help with concrete application reflects an incrementalist recognition that abstract behavioural scientific knowledge doesn't get automatically taken up within the broader policy system. Rather, it needs to be structured, concretized, and translated into more intuitive and action-oriented insights. The politico-strategic work performed backstage is also a shift toward incrementalism, understanding that the take-up of knowledge requires political persuasion and legitimization – in the absence of a universal consensus about what counts as valid evidence. To 'sell' their approach, behaviour experts harness the argumentative power of numbers, percentages and 'killer charts'. They construct clear-cut success-stories while making subtle choices that frame the outcomes in the most positive way (for an example, see the case of the behaviour experts at the City of Enschede in Appendix I). At the same time, it should be emphasized that this knowledge brokering role remains underpinned by a strong

rationalist ambition. It forms an incrementalist 'patch' in service of the further rationalization and scientization of the policy process. It is in this sense that two distinct institutional logics – rationalism and incrementalism – merge.

*Focusing on outputs*

The focus of the chapter will now turn to more internally oriented, substantive matters of behavioural policy practice – e.g. adopted theories, methods, quality criteria – and explore how they are adapted in light of competing logics. The balancing act discussed below specifically concerns how behaviour experts measure their own success. Rationalism would assume that social change can be 'crafted' by means of systematic policy analysis. Accordingly, it measures policy success by actual changes in policy outcomes (Lindblom 1959; Parsons 2002; Haynes et al. 2012). My observations however suggest that behaviour experts make a gradual shift toward incrementalism, in the sense that they moderate their outcome focus and also aim at making accomplishments earlier in the policy process. This is not to say that they fully operate according to a process-driven logic (Lindblom 1959). Rather, they balance between a process and outcome orientation. They adopt a mixed, modestly optimist attitude, on the one hand trusting their behaviour change capabilities and drawing confidence from dozens of successful examples, while also acknowledging the limitations to their efforts. They cannot guarantee successful policy outcomes. Rather, they can 'improve the chance of success' – through better underpinned analysis and design. Their realistically achievable success thus becomes one of 'engineering' the policy *process*, rather than actual outcomes.

Behaviour experts particularly derive their success from developing and refining policy analyses from a behavioural perspective. Behavioural analysis has a central place in the practice of Dutch behaviour experts (Feitsma 2019) and it involves, for instance, making step-by-step analyses of why citizens would invest in solar panels, how job seekers could be engaged in more compliant job seeking behaviour, and how CEO's could make more sustainable strategic choices. A core belief of behaviour experts is that behavioural knowledge helps to construct policy designs that are richer and better underpinned. While the refinement of policy analyses may yield improved outcomes in the end, in a more direct way it provides them a sense of argumentative legitimacy. With a clearer, thought-through narrative, it is easier to justify choices for particular policy designs. Table 7.4 evidences how behaviour experts authorize themselves by producing valuable outputs, revised policy analyses in particular.

**Table 7.4: Evidence of output focus**

| *Example* |
| --- |
| During an on-site conversation with a member of a ministerial behavioural team at the coffee-machine, he tells me a story that he used he used often to explain his team's vision and working method. His story comes from an episode from the TV-show South Park about 'Underpants Gnomes'. He describes how the episode features a group of gnomes with a 'genius' profit plan. They collect underpants and sell them, in three phases:<br>Phase 1: Collect underpants<br>Phase 2: ?<br>Phase 3: Make profit<br>Phase 1 is *what* they do, collecting underpants, and Phase 3 *why* they do it, making profit. Phase 2 concerns the reasoning behind how they will turn their collected underpants into profit. But the gnomes skip this phase. The behaviour expert explained that this is exactly what happens in policymaking too. Phase 1, the activities, and Phase 3, the expected results, are clear, but Phase 2, the reasoning about *how* your activities will lead to those results, is lacking. There is a 'world of behaviour' underneath a certain case that you first need to map out in order to formulate a clear rationale. The team helps to uncover that world and think policies through. That requires hard work and a structured approach, deciding on what behaviour is desired, which behavioural barriers and motives need to be triggered, and accordingly, which measures can do so. The expert tells me that Behavioural Insights is thus not only about effectiveness, but also about legitimacy. It is the underlying reasoning that matters. As a policymaker, you stand much stronger with a well-supported narrative, allowing you to explain why you do what you do. (On-site observation, 8 November 2016) |

*Analytical satisficing*

The previous section showed how behaviour experts, in the face of competing logics, do not focus strictly on actual behaviour change outcomes. Their realistically achievable success lies in developing valuable policy outputs like sounder policy analyses. This section follows on from this by looking further into how behaviour experts execute their analyses. It is argued that within this practice of policy analysis a balancing act is operative too. That is: while a rationalist logic assumes the possibility of and need for comprehensive analysis (Lindblom 1959), my observations suggest that behaviour experts make a gradual shift toward incrementalism as they relax their ambitions at comprehensive analysis. They distance their practice from a purely scientific logic, noting that it is 'not a university' and that 'the field is not a lab'. They also distance themselves from the need to produce complete analyses, knowing that in practice these will inevitably contain uncertainties and assumptions. Rather, they try to identify behavioural determinants and map out solutions and their

consequences as accurately as possible within their bounded capacity, while drawing confidence from the progress that *is* made.

The distancing from the need for full scientific certainty also reflects in how behaviour experts employ scientific models. While in the frontstage they may confidently present clear-cut models with which to analyse human behaviour, in the backstage they may point at the inconclusive nature of such models. They realize that such models are simplifying heuristics. An interviewee states: 'The power of models is that they are simplifications, abstractions of reality to make that reality more manageable'. Behaviour experts also point at the immense complexity of their cases in terms of why targets behave as they do. This 'richness of behaviour' is not easily mapped out, and doesn't per se fit in the scientific frameworks produced by behavioural scientists. This shift in how the status of these analytical models is perceived backstage, i.e. from conclusive frameworks to loosely guiding heuristics, reflects that experts do not expect their behavioural analyses to be wholly conclusive and certain.

Such *analytical satisficing* suggests that experts try to balance between rationalism and incrementalism. On the one hand, they are cautious not to take insufficient time for the analysis and 'jump to solutions' in the decision-making process. On the other, as timely, efficient, and widely applicable actions are needed, they are also cautious not to analyse too far, in too much detail, with too little tolerance for uncertainty and inconclusiveness. The balancing act is therefore one of gathering just enough facts, tolerating a degree of 'not-knowing', coming to crucial insights quickly without stranding in 'paralysis analysis'. Table 7.5 further evidences this balancing act.

**Table 7.5: Evidence of analytical satisficing**

| *Example* |
| --- |
| A behaviour expert talks about the workshop sessions he hosts, and the importance of striking a fine balance between analysing too little and too much: 'When we host a session and people find it interesting, you'll often see that they explore about two more ideas and that's it. You shouldn't bother them anymore with it then. That's the knowledge base. It's quite fragile. But we will go on in the process, over to the Design phase. Because we know that when you linger on in the Analysis phase, then people will drop out. It takes them too long. What they really want is to jump to solutions, but we're already restraining them from that. If we succeed in that, that's beneficial, but when you go on for too long, it becomes counterproductive. They'll say that it's going nowhere, so never mind. So you've got to push through. In a very big project we've experienced that you get commitment to do thorough analyses. But now in this (…) project, the participants only get a few homework assignments and we're already glad when they |

make those. It means that your analysis will improve, and that you're momentarily withholding people from jumping to solutions, which is very advantageous. But it's very relative. It depends on who you're working with, how much political-societal pressure rests on the case, and how much willingness there is amongst policymakers to invest in this. (…) So, yea, it's quite delicate. But everything you can achieve is an added benefit, is my feeling.'
Furthermore, interesting about the work method of this behaviour expert is the space that it leaves open for 'not-knowing'. In his workshops, he works with an analytical tool that explicitly asks users to list what is not yet known about a particular case. One of his team's core principles is: 'Check your assumptions: be aware that much of our thinking is based on assumptions. Make them explicit and make sure that you check them during the process.' (BIN NL 2017: 7).

*Horizontalizing the hierarchy of evidence*

The previous two sections showed behaviour experts balancing between an outcome versus processual focus, and between synoptic versus satisficed analysis. The balancing act discussed below deals with the type of methods that behaviour experts authorize as valid. A rationalist view starts from a clear 'hierarchy of evidence' in which knowledge gained from (meta-analyses of) RCTs gains the highest epistemic authority (Cairney 2017; Newman 2016). This view is however not wholly reflected in the field of behavioural expertise, in which RCTs appear to form a defining and yet also partially contested source of expertise. While some behaviour experts remain strictly tied to their 'what works' methodology, others challenge and depart from the rationalist idea that RCT methodology provides the 'the holy grail' of evaluation (also see Feitsma 2019). They point to the technical, methodological and juridical limitations of field experiments. Also, they note that other ways exist to gain valuable insights into the effectiveness of interventions (including qualitative and experience-based methods), thus arguing for the use of a wider palette of evaluation methods. Doing so, they move from a strictly rationalist hierarchy of evidence to a more pluralistic, *horizontal* version.

Those behaviour experts that move toward a more horizontal palette of evidence bases do not however devaluate RCTs entirely. They can be still directly involved in running and 'selling' experimental projects. But rather, they adopt a more nuanced stance, appreciating RCTs for their rigor and causal explanatory power, but simultaneously downplaying them in terms of their general usability. Specific concerns, *inter alia*, relate to: the inability to fully control field settings and thus conduct proper experiments; ethical and juridical arguments against randomizing treatment groups that thereby lead to unequal treatment of policy subjects; uncertainty about the transferability of experimental findings. Table 7.6 (first theme)

further evidences how behaviour experts partially distance themselves from a rationalist logic by deemphasizing the value of RCTs, arguing that they are 'just one tool in the kit of the researcher'.

**Table 7.6: Evidence of the horizontalization of the hierarchy of evidence**

| *Theme* | *Example* |
|---|---|
| Deemphasizing the value of RCTs | To improve waste sorting in high-rise buildings, one behaviour expert mentions that his team started a key project that followed a so-called 'royal route', moving from in-depth behavioural analysis to mapping out and *ex ante* evaluating prioritized solutions in the field with researchers running numerous RCTs. He however expresses some worry that despite the rigour of the 'royal route' in producing and evaluating interventions, in the end success would depend on how these interventions were locally implemented. A more nuanced stance on field experiments is expressed 'With the waste sorting project I'm a little bit afraid that per case it'll be so specific and context-dependent that you'd need to fully map out and understand the target group and the context in each and every case, before you can choose rightly between the possible intervention strategies (…). And that each subproject in every Dutch municipality will still become as extensive as [our project]. Except that (...) it does make it easier, the royal route. Because you know what matters and what you need to know to make head-way. That's already a whole lot… (…) and I think that the most important intervention strategies derive from following the royal route. That those strategies, in one way or another, do have much potential in specific situations. It's just that for the specific composition, the specific design, you'll still need to do your own analysis of all the particular aspects. And doing that is demanding, not something that I see happening in a lot of municipalities.' |
| Reemphasizing the value of non-experimental knowledge | A behaviour expert, working at the implementation agency Rijkswaterstaat, emphasizes the need for creative knowledge beyond insights gained from traditional behavioural analysis. Creativity is particularly necessary in making the step from analysing behaviours to designing fitting solutions. He keeps a list of all of his creative thinking techniques, varying from 'knowledge-based brainstorms' to 'I Ching' and the 'Synectics-technique', a methodology that works with metaphor and analogy to come up with original ideas. Besides creative knowledge, he also emphasizes the value of local knowledge brought in by regional public officials, but also local community |

> workers and citizens: 'How you make that translation [from analysis to local solutions]? That's quite a tricky one. I once organized a benchmark-session, just showing a movie of Cialdini and his principles [of social influence] and then letting the audience think about how they could apply this in [the policy area of] waste separation. You notice that they get through the movie quite easily and that applying each behavioural change technique to the specific situation and the identified "customer journeys" yields some nice results. So that's one nice way to do it. Those people also know the target group so they're able to think along. [It is important to] engage them with these techniques and let them think along about how to modify those or assess whether they work or not. That they develop a sensitivity for it. I don't think you can do all of that by yourself.'

Besides *de*emphasizing the epistemic authority of RCTs, a substantial group of behaviour experts also *re*emphasizes the value of nonexperimental types of knowledge in developing interventions and making informed statements about 'what works'. These experts work in ways that could be described as 'behaviourally-informed', loosely informed by behavioural science theories, rather than a 'behaviourally-tested' way that is strictly informed by local field experiments (Lourenço et al. 2016). As Table 7.6 (second theme, also see Appendix I) shows, this looser approach features an eclectic range of methods, including literature study, interviewing, and observing, but also common-sense reasoning, professional judgment, creative thinking, and brainstorming with local citizens. This suggests that behavioural policy is not necessarily coupled to a rationalist RCT-approach but can also be grounded in wholly different knowledge bases that are more in line with incrementalism.

## 7.5 Conclusion and discussion

A classic debate within public administration asks about the nature of the policy process. The academic consensus points toward an incrementalist nature (Bendor 2015). Yet, against the background of the flourishing of behavioural science, rationalism seems to have been revitalized in governments, exemplified by the rise of a 'Behavioural Insights' movement. This chapter has explored, based on ethnographic fieldwork in the context of Dutch government, how the members of this movement put their methods and ideas into practice in a not-so-rationalist environment. An important observation has been that, in pointing out where these behaviour experts fall on the rationalist–incrementalist spectrum, it differs whether you look at their frontstage or backstage operations. In the frontstage, behaviour experts push for

strong rationalism. However, in the backstage, they seek to harmonize their rationalist image with the incrementalism present in the wider policy system, resulting in various 'balancing acts'.

These findings challenge the idea of Behavioural Insights advocates as the 'new rationalists in town'. The real introduction of behavioural science in the existing policy system doesn't reflect the strong rationalism that is implicit in the dominant behavioural policy discourse (e.g. see Dolan et al. 2010; OECD 2017; Lourenço et al. 2016). It would be more accurate to depict that introduction as an ambiguous, negotiated and political process in which conflicting institutional logics merge and interact. Typically incrementalist features, like analytical satisficing, are mixed with typically rationalist features, such as the adherence to the rationalist policy cycle. The result is a hybrid, 'rationalized incrementalism' that operates in the backstage of behavioural public policy. This confirms earlier scholarly pleas advising not to approach the rationalism–incrementalism spectrum too binarily (Smith & May 1980; and in relation to the equivalent rationalism–constructivism debate: Newman 2016; Strassheim 2017). Rather than a dichotomy, a dynamic continuum exists in which rationalist and incrementalist forces coexist and blend on various dimensions (see Table 7.1). More fundamentally, this understanding questions the helpfulness of thinking in terms of two supposedly opposite and separated logics in the first place (Strassheim 2017).

The observation of this more hybrid, ambiguous logic operating in the behavioural policy process shows interesting analogies with the in and of itself ambiguous nature of the knowledge discipline by which behavioural policy is primarily underpinned (i.e. the new school of behavioural economics). An inherent ambiguity exists about new behavioural economics' stance toward rationality. It is on the one hand defined by its focus on how real people *deviate* from the rationality principle, yet it nonetheless still takes the rationality principle as its fundamental benchmark, thereby reproducing the idea of rationality (Sent 2004). Analogously, this chapter has shown how behaviour experts seek to cope with the ways in which real-world policymaking deviates from their rationalist ideals, yet while doing so they still take the wholly rationalized policy process as their benchmark and even come to defence of that benchmark in the frontstage. Thus, just as new behavioural economics relates ambiguously to rationality, so do behaviour experts, continuously switching between rationalist and non-rationalist assumptions across different organizational stages.

An interesting question is whether the observed compromising between conflicting institutional logics is a general feature of expertise being taken up in the policy system. This claim would be supported by other similar observations of how

institutional logics merge as new modes of policymaking make their appearance (e.g. Bjerregaard 2011). At the same time, the introduction of behavioural expertise may evoke a particularly strong clash between logics. The propagated hardcore rationalism of Behavioural Insights, with its strong assumptions about the readability, measurability and craftability of individual behaviour, stands in rather sharp contrast with the incrementalist logic underpinning real-world policymaking. Due to this stark clash, a greater need to compromise between logics can be expected.

The ability of behaviour experts to balance between institutional logics may well be a vital aspect of their legitimacy. Generally, as organizational practices are faced with competing demands, they will tend to search for ways that allow them to strike a tactful balance between those demands so that they are all sufficiently met (Brunsson 2007). This chapter suggests that behaviour experts are capably doing so too. They engage in particular balancing acts to negotiate between a competing rationalist and incrementalist logic. More broadly, they have constructed a rationalist frontstage which 'talks' differently, if not oppositely, in relation to what its backstage 'does'. For critical scholars, such inconsistency could be an easy opportunity to debunk yet another rationalist policy trend. However, from an organizational point of view, this inconsistency may exactly be what allows behaviour experts to thrive. It is a deliberate, *organized hypocrisy* (Brunsson 2007) that makes productive use of ambiguity. As behaviour experts compromise between competing logics and preach different messages across different stages, they show themselves to be tactically adaptable, able to meet seemingly incompatible demands at once.

At the same time, early signs of a less hypocritical, more self-reflective turn in the field of behavioural public policy may be observed. At the frontstage of the field, academic thought leaders have been exploring new questions, for instance looking into how elements of deliberative democracy can be incorporated into behavioural policy so that it becomes less technocratic (John 2018). Also, trendsetting units are devoting full reports to reflect on the limitations of a strong rationalist approach (Hallsworth et al. 2018). Such reflective developments seem to suggest that the critical behavioural policy literature is being heard, and that also on a broader level, debates on the limitations of evidence-based policymaking as well as discussions about the 'replication crises' in the fields of medicine and psychology (Deaton & Cartwright 2018) are impacting the field, also in its frontstage discussions and agenda setting. The front- and backstage of behavioural public policy might thus be coming closer together.

*Chapter 8*

# Deliberating the Behavioural State

*A 'Behavioural Insights' movement has emerged within governments. This movement infuses policymaking with behavioural scientific insights into the rationally bounded nature of human behaviour, hoping to make more effective and cost-efficient policies without being too obtrusive. Alongside sustained admirations of some, others see in Behavioural Insights the threatening revival of technocracy, and more particularly a 'psychocracy': a mode of public decision-making that wrongfully reduces the world of policymaking to a rational-instrumental and top-down affair dictated by psychological expertise. This chapter argues, however, that the claims of technocracy and psychocracy are overgeneralizations, emanating from a frontstage-focused debate that ignores a vast backwater of emerging behavioural policy practices. Grounded in four case studies on behavioural policymaking in Dutch governance, it will be demonstrated that at least part of this backwater is neither so technocratic nor so psychocratic as the critics claim.*

---

Chapter 8 is published as a discussion and commentary article by Joram Feitsma in Policy Sciences.

Feitsma, J.N.P. (2018). The behavioural state: critical observations on technocracy and psychocracy. *Policy Sciences, 51*(3): 387-410. doi: 10.1007/s11077-018-9325-5

## 8.1 Introduction

A 'Behavioural Insights' movement has arisen within governments in the last decade. While this movement is still 'under construction', it has already become the object of a vast polemic debate, with some critics unmasking it as an updated form of technocracy with a psychological twist – i.e. a 'psychocracy'. This chapter, however, argues that the portrayal of Behavioural Insights as technocratic and psychocratic is founded on only a handful of emblematic success stories which are not wholly representative for some of the alternative practices that have emerged in the 'backwater' of the Behavioural Insights field. I substantiate this argument with four case studies of not-so-technocratic and not-so-psychocratic behavioural practices, grounded in ethnographic fieldwork on upcoming behaviour experts in Dutch government.

The Behavioural Insights movement is part of a global policy agenda expressing a renewed interest in insights and methods from the behavioural sciences when underpinning and forming public policies (e.g. Strassheim & Korinek 2015; John 2018; Whitehead et al. 2017; Lunn 2012; Lourenço et al. 2016). An overarching behavioural insight is the idea that humans are not merely rational beings but have forgetful, lazy, ignorant, impulsive, and other less rational traits that make them behave against their own interests or goals. As the behavioural economics bestseller *Nudge* puts it: humans are more like Homer Simpson than *homo economicus* (Thaler & Sunstein 2008). The increased awareness hereof has resulted in the development of a more psychological style of governance that seeks to change behaviour by subtly redesigning local environments in ways that try to acknowledge and tap into the bounded rationality of policy subjects – all this with the support of rigorous experimental evaluation methods. This new style reflects in a changing professional apparatus of governments. Governments of today now inhabit 'Choice Architects', 'Nudge Experts', and 'Chief Behavioural Officers' (John 2018; Feitsma 2019).

Perhaps the most vivid manifestation is the launch of 'Behavioural Insight Teams' (BITs): specialized units that are trialling novel types of 'nudge' interventions in a broad range of policy domains, from public health to tax compliance to sustainability. By now, there are BITs all across the globe, both inside and outside of government, from a 'Behavioural Economics Team of the Australian Government' to a 'Qatar Behavioural Insights Unit' to a 'Behavioural Insights Network Netherlands' (BIN NL). The initial trendsetter, the BIT in the UK, has grown from less than 10 to over 100 employees, with offices in London, New York, and Sydney, and has run over 600 Randomized Controlled Trials (RCTs). It has helped with, *inter alia,* promoting diversity in the police-corps, helping people to find a job faster, and getting people to

pay their taxes on time (Dolan et al. 2010; Service et al. 2014). In the area of tax policy, for instance, it famously ran an experiment that increased tax return compliance with 5.1%, collecting an extra of GBP 9 million during a 23-day period, simply by including in the letters to late tax payers that 'Nine out of ten people pay their tax on time. You are currently in the very small minority of people who have not paid us yet.' (OECD 2017: 342). The British BIT has been one of the key missionaries in service of the 'behavioural turn' within governments, carefully constructing and promoting the Behavioural Insights 'brand' via a range of papers (e.g. Haynes et al. 2012), update reports (e.g. BIN NL 2017a) and books (Halpern 2015). Besides these specialized behavioural units, the behavioural turn is also boosted by the received support of distinguished politicians (e.g. David Cameron and Barack Obama), as well as the promoting work of influential transnational policy actors (e.g. the OECD, the World Bank, and the European Commission). Furthermore, behavioural scientists have increasingly been offering their expertise directly to governments. For instance, Richard Thaler, co-author of *Nudge* and recent Nobel Memorial Prize laureate, has been one of the key players in getting behavioural insights on the British policy agenda. Taken together, the new BITs and behaviour change professionals, along with their supporting top civil servants, think tanks, politicians, and academics, form a powerful frontstage chorus that speaks very positively about the emerging behavioural state, firmly believing that behavioural insights will help to make policies more effective and cost-efficient while still respectful of liberal freedoms (Dolan et al. 2010; Thaler & Sunstein 2008; World Bank 2015).

However, this celebratory if not self-advertising (Campbell 2017) chorus represents only one voice in the behavioural policy debate. Right from the beginning, and again particularly in the Anglo-Saxon sphere, also scepticism and criticism has been expressed against Behavioural Insights whether that be in public media, political or academic discourse. In the public media, the British BIT was welcomed with headliners as 'Nudge Nudge, Say No More. Brits' Minds Will Be Controlled Without Us Knowing It' (Hunt 5 February 2014), and *Nudge* co-author Cass Sunstein was nominated as 'the most dangerous man in America' (Beck 28 September 2010). In politics, the House of Lords expressed concern about the claimed added value of the British BIT, leading to an official investigation into the performance of this unit (House of Lords Science and Technology Select Committee 2011). In academia, many scholars have voiced critiques too. They have been sceptic about the added benefits of behavioural insights to policymaking and have also voiced concern about the underlying political and moral agenda of Behavioural Insights. So far, a range of normative critiques on the emerging behavioural state exist (Jones et al. 2013; Furedi 2011; McLaughlin 2016; White 2013; Whitehead et al. 2017; St. Paul 2011; Mulderrig 2017a, b; Mettler 2011; Leggett 2014; Rowson 2011; Campbell 2017). Initial

concerns foresee the loss of liberal freedoms and autonomy, questioning whether citizens can direct their lives according to their own values when their everyday environments are increasingly supplemented with subtle governmental interventions. Further uneasiness is expressed with the idea of bureaucratic experts deciding what is good for citizens, possibly overriding their preferences (e.g. White 2013), leading particularly to the marginalization of lower-class income groups (Mulderrig 2017b). Underneath such concerns seems to lie a more generic concern about the possible upsurge of *technocracy*. By technocracy I mean a model of public decision-making in which bureaucrats decide *for* rather than *with* citizens, guided by scientific expertise rather than political dialogue (Clarence 2002). This 'technocracy claim' is partially due to that the nature of behavioural policy can be 'stealthy' (Whitehead et al. 2017; Mols et al. 2015). It can target and operate through the unconscious, automatic, *System I* (Kahneman 2011) decision-making faculties of citizens, and can take the form of pilot experiments that take place outside the public eye. This can make it rather difficult for citizens to notice or contest behavioural policies, let alone get involved in the decision-making process.

Disquietude over the resurgence of technocracy as the result of policy innovations is not new. The emergence of the policy analysis movement after World War II raised similar concerns (Fischer et al. 2015), just as, more recently, the rise of evidence-based policy (e.g. Cabinet Office 1999) was explained as 'technocracy reinvented' (Clarence 2002). However, it should be noted that Behavioural Insights is associated with a particular technocratic form: a *psychocracy* (Jones et al. 2013). Psychocracy is characterized by its exclusive use of psychological knowledge and methods in the governance of citizen behaviours. This psychocratic nature of the behavioural state has also been criticized specifically (Mulderrig 2017a; Whitehead et al. 2017; Mols et al. 2015; Rowson 2011; Jones & Whitehead 2018; Feitsma 2019). One problem is that it has an overly psychologized idea of people's 'environments', only looking at immediate, physical and technical aspects at the micro-level that have proved to affect decision-making in psychological experiments (Whitehead et al. 2017; Rowson 2011). This neglects that the environment is also shaped by deeper sociocultural, institutional, political, and economic forces at the macro-level that are much less 'craftable'. Furthermore, criticisms have been voiced regarding the bias of the behavioural state towards a particular type of evidence, i.e. experimental knowledge (Feitsma 2019). Although such knowledge provides important information about 'what worked there-and-then' (Biesta 2007), policymaking would benefit from, if not necessitate inclusion of other types of valuable evidence (e.g. qualitative or experiential knowledge) in the policy process (Parsons 2002; Scott 1998; Fleming & Rhodes 2018). In sum, the 'psychocracy claim' holds that Behavioural Insights reduces the complex world of policymaking to a simplistic, monodisciplinary, and

desocialized affair in which there is only space for the ideas and methods of psychological experts.

As its central empirical contribution, this chapter seeks to problematize the above-mentioned claims about the upsurge of technocracy in general – and psychocracy in particular – that are implied by the current behavioural policymaking trend. These claims are problematic because they emanate from a cemented 'trench warfare' type of debate in which participants speak only to and for their own 'trenches' (for an exception, see John 2017). This results in echo chambers that obstruct actual and nuanced debate. Moreover, the debate is muddled because its participants tend to make their arguments merely based on the frontstage manifestations and role models of Behavioural Insights. Antagonists have partly formulated their critiques based on abstract (dystopian) sketches about what a 'nudge-world' (Waldron 9 October 2014) might look like or on the iconic idea of behavioural policy as constructed by role models such as the British BIT, *Nudge*, and other published works (e.g. Dolan et al. 2010; Haynes et al. 2012; Service et al. 2014; Halpern 2015). Because such role models tend to present generic ideas about how behavioural policy *should* be developed rather than how it actually *is* practised across various domains, these ideas do not necessarily offer an accurate and complete empirical basis for critical inspection.

In fact, in the debate's focus on the frontstage facets of Behavioural Insights, a vast backwater of emerging behavioural practices is ignored. The Dutch government, for instance, is part of a glossed over hinterland, where nonetheless manifold behavioural practices have blossomed in the last decade (e.g. Schillemans & De Vries 2016b). While there are some more substantial networks, such as BIN NL at national level and Behavioural Insights Group Rotterdam at municipal level, most of these practices are relatively small, informal, and in an experimental phase (Feitsma & Schillemans 2019). With the disregard of this backwater, the risk looms that an imprecise if not misrepresentative idea of behavioural policy is formed. To avoid this, this chapter sheds a new light on the behavioural policy debate – particularly with regard to the alleged upswing of technocracy – by shifting the current focus on the frontstage of Behavioural Insights over to some of the lesser-known backstage developments. I will zoom in on four different cases of behavioural policy practice, based on ethnographic fieldwork on behaviour experts in the Dutch government over a period of 4 years. By investigating these practices in relation to the themes of technocracy and psychocracy, I seek to stimulate critical thinking about the emerging behavioural state and expand it beyond the more well-known concerns about hampered autonomy and deepened paternalism (e.g. McLaughlin 2016; Jones et al. 2013). The key point, briefly put, is that the technocracy and psychocracy claims are overgeneralizations. The presented case

studies show that behavioural practices can actually be quite democratic and inclusive of non-psychological expertise. While this does in no way make the critical claims irrelevant, it does suggest that they are at least unreflective of some developments in the field and the promise of a more 'deliberative' variant of behavioural public policy.

## 8.2 Methods

### *Design*

This chapter takes off from an anthropological perspective that views the behavioural state not as a uniform, coherent, and abstract entity but rather as an assemblage of different, competing, and contradictory practices that operate both within but also further out of the deep state (Jones et al. 2013). From this perspective, the behavioural state is a peopled state, and it is these people that should be carefully studied. It is worthwhile to do 'up close and personal' ethnographic type of fieldwork (Rhodes et al. 2007) on these behaviour experts and their practices, to gain a deeper understanding of the behavioural state including its better hidden backstage realities (Van Hulst 2008). While a small body of qualitative studies on the behavioural state exists (e.g. John 2014; Jones et al. 2013; Whitehead et al. 2017), these studies remain largely grounded in generic and 'thin' empirical data, paying little attention to everyday experiences in the backstage.

### *Data collection*

This chapter is part of a larger ethnographic study on the rise of behaviour experts in government. For this study, I have interviewed, observed, and collaborated with behaviour experts in various levels of Dutch government for over a period of four years. The research has been undertaken in four research phases:

- From 2014 to 2016:
    - Interviews with 35 central governmental behaviour experts
    - 55 hours of short-term participant observation at various sites
    - Involvement in a local 'Urban Nudging' project over the course of 10 months
    - Document study
- In 2016:
    - 4 months of long-term participant observation as employee ethnographer in a small ministerial BIT, plus several pre- and post-visits to the field
    - Document study
- From 2016 to 2018:
    - Interviews with 15 local governmental behaviour experts

    - 19 hours of short-term participant observation at various sites of local governance, primarily municipalities
    - Document study
- In 2018:
    - Focus group with a mixed group of governmental behaviour experts

I started the research process with a mapping exercise of behaviour experts in Dutch government, based on an initial document study and the use of the snowballing technique. My sample only included self-identifying governmental behaviour experts who explicitly claimed to be making use of behavioural insights on a structural basis. My sampling strategy was partly pragmatic, doing participant observations wherever I got access, but also driven by the aim to capture the field of Dutch behavioural expertise in its full breadth. I ensured that my sample represented a broad range of public organizations, from ministerial departments to regulatory and executive agencies. Furthermore, I later conducted an extra round of ethnographic fieldwork, this time specifically zooming in on behaviour experts in local governments – which had up to then been a blind spot in my sample.

In terms of the conducted fieldwork, the first and third round of data gathering can best be described as a 'yoyo' type of fieldwork (Wulff 2002), making short visits to a broad range of behaviour experts in their professional habitats. Here, the focus was on doing semi-structured interviews. These were held in a closed setting, mostly where the interviewees worked, and guided by a protocol asking about general features of their work (e.g. organizational design, goals, tasks, everyday routines, successes, challenges) yet also leaving ample space to follow their own leads on where to take the conversation. The interviews were recorded (when having received permission to do so), selectively transcribed and turned into field reports. Next to interviews, I conducted various, mostly short-term participant-observations – such as shadowing behaviour experts for a day, attending internal work meetings, visiting educational meetings, having on-site informal conversations, and directly collaborating in projects. During or right after these observations, I jotted down quick notes, which I then translated into more elaborate field reports. The 'yoyo' type of approach initially helped to study the field of behaviour experts in all of its diversity. To achieve a deeper immersion in the field and see more of its backstage, I also collaborated more directly with behaviour experts, getting involved as an academic adviser in a local project that sought to incorporate behavioural insights. Moreover, as part of a second round of fieldwork, I was seconded to a Dutch BIT as an employee-ethnographer. Taking on the role of behaviour expert myself allowed me to develop a more tacit and experience-based understanding of behavioural policy practice. This whole trajectory of intensive and immersive fieldwork – going beyond the

organizational frontstage, talking to and shadowing a multitude of experts, across various loci, over the course of years, while triangulating methods (i.e. interviews, participant-observation, document study) and organizing peer feedback – hopefully serves to bolster the trustworthiness and plausibility of my 'conjectures' (Rhodes 2014).

### *Data analysis*

The analytical process of my overarching study on behaviour experts also has a 'yoyo' element in it, constantly moving back and forth between doing fieldwork, 'headwork' and 'textwork' (Beuving & De Vries 2014; Bowen 2008). My earlier research projects were oriented on *observing* and *describing* actual practices (see Feitsma 2019), professionalisms (see Feitsma & Schillemans 2019), and rationalities of behavioural policymaking. As the research progressed, and the iterative interplay between data collection and data analysis continued, my agenda also shifted to *debating* behavioural public policy. One particularly emerging topic for debate was that there seemed to be tensions between the rather polarized *Nudge* debate with its Anglo-Saxon focus and the lesser-known behavioural policy practices that I was studying. I sensed some discrepancy between what was happening within this backwater of behavioural public policy I was studying and how it was portrayed by some prominent *Nudge* critiques, claiming that behavioural policies are universally illegitimate, overpromising and ill-informed (e.g. Furedi 2011; Mols et al. 2015; St. Paul 2011). This chapter arose from this discrepancy. To do so, the analytical process for this particular study has been as follows. I started by breaking down the Behavioural Insights critiques based on a literature study. I then compared these with my fieldwork observations, and selected the two critical themes that most saliently seemed to point at a misfit between observed backwater practice and radical critique. I further worked out these themes, and their underlying assumptions (see Table 8.1 in the conclusion and discussion), and took them as sensitizing concepts to guide further empirical scrutiny. I then rescreened my data in light of these themes, looking for specific cases that made salient the misfit between the critical claims and my field observations. This screening followed a purposive selection strategy, searching for 'information-rich cases' (Bowen 2008) that best spoke to my sensitizing concepts and of which I had collected rich data from multiple sources – so as to increase trustworthiness. It needs to be emphasized that the case selection was based on the specific theory-driven question of whether counterexamples existed to the radicalizing technocracy and psychocracy critiques. I deliberately looked for counter cases, and my sampling strategy herein followed a logic of achieving theoretical adequacy – not a logic of achieving empirical representativeness. This means that the representativeness of my case studies should not be overstated (for further discussion on this, see the Discussion section). Nevertheless, putting the spotlights on such

counter cases is important, given the overwhelming general academic consensus that behavioural public policy in its current stage of development indeed has a technocratic and psychocratic nature (John 2018). It may help to arrive at a more realistic, differentiated understanding of behavioural public policy.

The core analysis that follows consists of an analysis of four cases of lesser-known practices that take place in the backwater of Behavioural Insights. First, I explore two cases in relation to technocracy, and then, more specifically, two cases in relation to psychocracy. The case studies are introduced with a recap of the theoretical debate around these claims.

## 8.3 Technocracy

The general critical claim investigated in this chapter concerns technocracy: a command-and-control model of public decision-making that is reliant on scientific expertise rather than political debate (Clarence 2002). A first sign of the technocratic character of Behavioural Insights could be derived from its basic ambition to infuse policymaking with (new) insights from the behavioural sciences so as to make policies more effective. This ambition thus combines a focus on the scientization of the policy process with an aim to optimize the effectiveness of policies. Implicitly, this shifts the focus away from more democratizing lines of inquiry, such as investigating current political needs or exploring how citizens can be involved, what knowledge they might have to offer, and what their preferences are.

The concern about the exclusion of citizens from the policy process particularly arises from the alleged 'stealthy' nature of behavioural policies (Whitehead et al. 2017; Mols et al. 2015; Mettler 2011). This stealthiness relates to the observation that behavioural policy tends to take shape in the form of rather subtle interventions, which in some cases deliberately target the unconscious decision-making faculties of citizens (although besides 'exploitative' there are also more 'educative' types of nudges, see Schubert 2017; Sunstein 2017a), and which sometimes affect citizens as part of undisclosed experiments. For instance, various Dutch government agencies have been making subtle changes in the letters that they send to policy subjects, inspired by the work of BIT UK in the context of tax policy (Service et al. 2014). Apart from the 'Team Behaviour Change' of the Dutch Tax and Customs Administration, which shared some of the results of its nudge-experiments in the public media (De Jong & Rusman 2 March 2015), there is little public information available about the exact nudge-interventions that these agencies have been using, nor about the letter experiments that they have been conducting. These stealthy aspects of behavioural policies make it difficult for citizens to observe, debate, reject, or give their informed

consent to them. They hinder the formation of Habermasian 'deliberative spaces' in which behavioural governments are held accountable for their actions (Whitehead et al. 2017; McLaughlin 2016). Put differently, such developments may contribute to the rise of *the submerged state* (Mettler 2011), in which government actions increasingly become concealed to the general public, thereby hampering the possibilities of democratic decision-making and obfuscating the relationship between governments and citizens.

And yet, if we look at the actual behavioural policy practices that have emerged throughout governments, the portraiture of Behavioural Insights as technocratic seems somewhat problematic. I will mention some broad and preliminary observations. To begin with, ample attempts have been made to be transparent and create deliberative space about behavioural policy, for instance through various BIT update reports (e.g. BIN NL 2017a), books (Halpern 2015), and scientific articles (John 2014). In the particular Dutch case, behavioural policymaking has also been subject to formal deliberation, with government advisory councils and behaviour experts organizing debates and writing reports about behavioural policy; one could even say there is more deliberative space than operational activity regarding behavioural insights in the Netherlands (e.g. see BIN NL 2017a; Jonkers & Tiemeijer 2014; Van Staveren et al. 2014). The technocracy claim is furthermore challenged by the observation that the *targets* of behavioural policy are not per se citizens. In my interviews and participant-observations, I have found various cases in which other target groups were subjected to behaviour change tactics, such as commercial enterprises or policymakers. Moreover, it is also the case that the *makers* of behavioural policy are not per se governments. More often, these policies come about in wider fields of governance, in which governments collaborate with many other societal partners, playing a rather distanced – *meta-governing* (Sørensen & Torfing 2009) – role in the actual development and implementation of behaviour change interventions (Feitsma 2019). The observation that behavioural policies are neither per se made *for* citizens nor *by* government officials doesn't align with a technocratic frame of the state managing the citizen.

To give further empirical weight to the argument that the generalized technocracy claim is problematic, I will zoom in on two lesser-known cases of behavioural policy practice. These practices might even be considered to be democratizing, given their attention for the empowerment of citizens and public deliberation about new policymaking styles.

***Urban Nudging***

*Background characteristics*
- Temporary project at local level
- Launched in 2015
- Collaboration between academics, artists, creative designers, and municipal officials
- Goal: facilitate societal debate about nudging

The first case-study features a rather atypical, small scale, and local project, called 'Urban Nudging', which was set up in the spring of 2015 in the Dutch city of Utrecht. This project was based on a collaboration between creative designers, artists, academics (including myself), and municipal public officials. The underlying goal of the project was to illustrate how behaviour change strategies could be developed in practice, so as to increase citizens' own competencies in behaviour change. The specific challenge that Urban Nudging tackled was the wrong parking of bicycles in the city centre. This was a problem that the City of Utrecht had been struggling with for some time, as there were several 'hotspots' in the city centre where wrongly parked bicycles caused significant hindrance for pedestrians. After a short introduction into behavioural design and a field exploration, three different groups of artists and designers were deployed to design and realize their own nudges for the bike parking problem in the city centre, under the guidance of the project team and academic experts. The resulting interventions (see Figure 8.1) were diverse in terms of their forms and the behavioural mechanisms that they sought to address. The more playful ones included a road sign warning bikers to watch out for a 'bike monster eating up wrongly parked bikes'. A more sophisticated behavioural design was thought of by a group of landscape designers who analysed the various walking routes of pedestrians on a particular street, and then explored ways to visualize those routes, for instance by spraying traffic lines on the street, as a way of making bikers aware of other road users.

After having devised and implemented its behaviour change strategies for bike parking, Urban Nudging communicated its results to a broader audience of citizens in a project report (SETUP 2015). It developed an 'Urban Nudging Kit' that included the 'ingredients' of some of the developed nudges, meant to enthuse citizens to employ nudging techniques for themselves.

**Figure 8.1: Urban Nudges**


The lead in Urban Nudging was taken by SETUP, a small enterprise that focuses on societal innovations. SETUP launched this project out of an awareness of the popularity of novel behaviour change strategies. While it recognized the societal potential of these developments, it also noticed that these insights and behaviour change techniques were mostly taken up by businesses and governments. Citizens remained largely unaware of them. Believing that 'knowledge about nudging should not only be reserved to a limited selection of people in power' (SETUP 2015: 9), SETUP's aim was thus to, in its own phrasing, 'democratize' nudging by introducing citizens into the world of behavioural insights:

> In a series of playful nudges, we introduced the phenomenon in analogous forms to citizens of Utrecht. By stimulating them to make their own nudges, we have opened the debate about practices that are already influencing us on a daily basis, both offline and online. Urban Nudging is a democratization of the nudge; we celebrate the phenomenon of nudging, make these tricks

> accessible to everybody and activate citizens to think about the consequences, possibilities, and lastly the desirability of these practices. (SETUP 2015: 7–8)

This quote reveals an ambition to stimulate public deliberation about the use of behavioural insights. Urban Nudging tried to do so in a distinctive and interesting way. It sought not only to make fellow citizens aware of how they were continuously affected by public and private behaviour change strategies, but also aimed to facilitate them in using these strategies themselves. Hence, the project could be interpreted as empowering citizens in two distinct ways, both at odds with the technocracy critique. First, it aimed to increase their *behavioural literacy* (Whitehead et al. 2017) by making them more aware of everyday behaviour change efforts and the psychological mechanisms at work in these efforts. Second, the project aimed to develop the *behaviour change capacity* of citizens by showing how they themselves could harness the power of behavioural insights. This would enable them both to counteract behavioural influences imposed by external societal actors and also to 'become part of an actual bottom-up problem-solving capital' (SETUP 2015: 9).

Urban Nudging can be understood as an attempt to hold off a technocratic variant of behavioural policy. This case challenges the technocracy claim, first, by rejecting the assumption that behavioural policies are per se orchestrated by state officials. In Urban Nudging, state officials played only a marginal role. The lead was instead taken by a cultural enterprise which linked together actors from academia and the artistic domain. Furthermore, the idea that behavioural policies are accompanied by a lack of public deliberation and co-production is called into question. The central aim of Urban Nudging was to involve citizens in the behaviour change arena. Its efforts to build a decentred, bottom-up, collaborative behaviour change capital belie the claim that behavioural policies by necessity are produced and delivered in a command-and-control fashion.

### *Litter-free communities*

*Background characteristics*
- Structural, centrally orchestrated project
- Launched in 2017 but with earlier roots
- Broad, multi-level partnership with various government agencies and civil society actors
- Goal: develop effective anti-litter policy

While the previous case study featured a temporary project situated at the local level, in what follows a more structural project will be discussed, orchestrated by central government policymakers. In 2017, Rijkswaterstaat, the executive agency of the Dutch

Ministry of Infrastructure and Water Management, collaborated with civil society actors to promote a national policy plan to reduce littering, called the 'Landelijke Aanpak Zwerfafval (LAZ)' [National Plan on Littering]. As an interviewed program coordinator of LAZ noted, this plan relied strongly on behavioural insights, particularly when it came to analysing which behavioural factors were at stake regarding the issue of littering. The program coordinator emphasized the importance of looking at unconscious and emotional aspects of decision-making and recognizing what he called 'the power of the context' in shaping human behaviour. *Prima facie,* LAZ could be deemed technocratic as in essence it sought to analyse and manage citizens from a distance with the help of science. This basic technocratic ambition reflected in the development of various anti-littering strategies on the basis of new behavioural insights (e.g. Novi Mores 2017).

What makes LAZ less technocratic, however, is its focus on participation and citizen activation. Based on the behavioural mechanisms behind perceived ownership and peer-to-peer influence, LAZ was underpinned by the belief that involving citizens in the decision-making process would yield more effective results. The perceived importance of citizen participation is exemplified by a collaboration of LAZ with a renowned behaviour change consultancy. This consultancy produced an advisory report on 'Stimulating and maintaining participation regarding littering' (Dijksterhuis & Van Baaren 2015), strategizing about how citizen participation could be facilitated, how participants could be kept 'happy and active', and also how municipalities should deal with initiatives from citizens. The report featured a menu with several strategies for municipalities to activate citizens. These strategies were thoroughly grounded in behavioural insights, suggesting subtle contextual readjustments that paid attention to the importance of symbolism, presentation, social norms, wording, and the framing and ordering of information. Suggestions included: making use of key active figures in the neighbourhood; using natural contact moments (e.g. local festivals) to communicate the desired behaviour; harnessing the power of reciprocity by handing out helpful tools (e.g. gripper waste pickup tools) as gifts; using communication signs that prime citizens with the idea of 'togetherness'; turning communal cleaning events into pleasant experiences; using social rewards (e.g. sending out Christmas cards, thanking citizens for their contributions in keeping the community litter-free); and using the so-called 'foot-in-the-door' technique, first getting citizens to make a small commitment, after which to ask for a larger commitment which citizens will then be more likely to make for the sake of being consistent (Dijksterhuis & Van Baaren 2015).

The fact that LAZ sought to stimulate participation by hiring a consultancy firm specialized in behaviour change interestingly reflects the fusion of two seemingly

conflicting policy strategies: the individual behaviour change focused 'Nudge' with the collective deliberation and participation oriented 'Think' (John et al. 2009). In LAZ, participation was viewed as a way of evoking behaviour change, and behaviour change strategies were used to stimulate and maintain participation. This interaction between 'Nudge' and 'Think' casts doubt upon the depiction of behavioural polices as purely technocratic. Rather, the less deliberative 'Nudge' tactics used in LAZ were used as second-order strategies in service of carrying out first-order strategies that sought to involve and activate citizens. On this first-order level, behavioural policies thus served a facilitating rather than impeding role in the cocreation of policies.

The fusion between 'Nudge' and 'Think' nevertheless raises questions about the assumed voluntariness and autonomy behind participation projects, and the desirability of governments undertaking concerted efforts to 'nudge citizens into participation'. The LAZ case suggests that citizens' choices to participate in public decision-making are not per se made wholly autonomously, but instead can be the result of a deliberately shaped process. This observation ties well with a Foucauldian perspective showing how late modern governments exercise power over citizens *through* freedom and self-regulation (Jones et al. 2013). In the LAZ case, what is at the surface perceived as citizens freely choosing to increase their control over their own communities, is actually impacted by subtle manifestations of governmental power that seek to orchestrate citizen behaviours through activation practices.

Besides its endeavours to involve citizens in the development of anti-litter policy, LAZ also sought to activate local government officials. To illustrate, an interviewed program coordinator noted that a substantial part of his work consisted in changing the behaviour of his peers in local governments, trying to involve them in the domains of littering and recycling policy. He framed his own challenges as follows: how could he get the issue of littering and recycling on top public servants' policy agendas? And how might he persuade municipal officials to start using behavioural insights techniques within these policy domains? Recognizing the important role of municipal officials in realizing behaviour change, he sought to empower them with special guidelines, tools, training sessions, and long-term advisory trajectories educating them about behavioural insights. He also sought to motivate them using behaviourally informed strategies such as informing municipalities about their performance in relation to both the nationwide goals and the performance of other municipalities (i.e. giving feedback and communicating social norms). Also, he pressed top-level officials to write formal pacts as a sign of their engagement with the issue (i.e. leveraging the power of public commitment). A more sophisticated strategy was the development of a 'Serious Litter Game'. This game presented involved policy

actors with a fictional case in which they had to make policy on littering while taking on different stakeholder roles. This act of gamification was meant to put littering on the policy agenda in a 'nice, relaxed, and effective way', to stimulate an interactive and deeper learning process, and to have participants empathize with each other's perspectives and 'develop intuition about litter policy' (Gemeente Schoon 2015). As these above-mentioned examples, and also the following quote, suggest, the program coordinator thus engaged with behavioural insights on a meta-perspective, seeking to change the behaviour of an intermediary group of policy officers who would on their turn then be better equipped to change the behaviour of citizens regarding littering and recycling:

> How do I get the municipality to engage with this [policy]? That's the first line of behaviour change that you've got in mind. So you've got to apply your knowledge also to persuade this target group. And the second line is: you want them to do it right, with a methodology, the right measures, the right consultancies, the right guidance. So we organize training programs and do benchmarking. (…) For instance, also in the program of household waste with regard to the role of municipalities, we applied A–M–O [ability–motivation–opportunity]. First we've got to motivate them and then we've got to give them the opportunities. If we can motivate management, resources will be made available, so then officials will get the opportunities to get started. In-between that lies ability: how do you make a proper policy proposal? Which measures do you use? How do you deal with residents? These sorts of things.
> *So you're also nudging local policy officers?*
> Well, of course. I mean… you want to convince them… persuade them. How do you do that? There's obviously a range of techniques that you can employ for that.

The focus in LAZ on activating local officials can be interpreted as another instance in which behaviour change strategies are fused with participatory aims. This particular case furthermore demonstrates that not only *citizens* but also *policymakers* are being subjected to behaviour change policies. This problematizes the technocracy claim which views the rise of Behavioural Insights in a rather one-sided way of how the state impacts its citizens, neglecting the behaviour change dynamics between the state centre (i.e. national government) and its peripheries (i.e. municipal governments) as illustrated in LAZ. When it is the state itself that is being affected by the behaviour change agenda, with behaviour experts practicing what they preach in their own policy networks, concern about the technocratic exclusion of citizens from the policy process seems less relevant.

## 8.4 Psychocracy

From exploring the generic technocracy claim, the chapter now moves to investigating the more specific claim concerning *psychocracy:* a technocratic form that is specifically and exclusively grounded in psychology (Jones et al. 2013). This dominance of psychology relates both to the *ideas* that are included and the *methods* that are authorized as valid. With regard to both of these aspects, Behavioural Insights has been criticized (Mulderrig 2017a; Whitehead et al. 2017; Jones and Whitehead 2018; Feitsma 2019; Mols et al. 2015). On the level of ideas, the psychocracy claim holds that Behavioural Insights overly relies on psychological theories. Moreover, within this domain, it tends to focus exclusively on cognitive psychology, with less attention for the more social, motivational, and psychodynamic branches. A result of this is that the fields understand the 'environment' in a narrow, overly psychologized way (Whitehead et al. 2017). It thinks it can realize behaviour change simply by readjusting aspects of the immediate, physical, and technical environments of citizens – the so-called 'choice architectures' (Thaler and Sunstein 2008). Such an approach however fails to acknowledge the 'obstinate nature of environmental legacies and the agency of the material world' (Whitehead et al. 2017: 158). Environments are not as readily mouldable as Behavioural Insights claims. Moreover, environments do not simply consist of material objects, but rather of a complex amalgam of both material and immaterial objects shaped by micro-level effects but also by wider political, economic, sociocultural, and institutional forces (Shove et al. 2012). The disregard for such wider macro-level forces leads Behavioural Insights to believe that behaviour change can be 'crafted' through mere technical choice architecture. For instance, in the area of obesity policy, it may satisfice with subtly readjusting school canteen settings, rather than seeking to change the systemic economic inequalities that underpin unhealthy eating behaviours (Mulderrig 2017a).

On the level of methods, the psychocracy claim points at Behavioural Insights' problematic understanding of 'evidence'. Behavioural Insights' underlying *evidence-based policy* ideology has been criticized for its rather peculiar understanding of the policy process and the role of evidence therein (Parsons 2002; Ingold & Monaghan 2016; Cairney 2017; Feitsma 2019). To begin with, Behavioural Insights thinks in a rather instrumental way about science-policy relation, as if academic knowledge flows naturally to the world policy, unaffected by political motives and interests. Such rational-instrumentalism however overlooks various by now well-documented observations of bounded rationality within the policy system, such as that policymakers tend to engage in satisficed decision-making based on incomplete analyses, dislike novel and complex, and maintain the status quo or change it only incrementally (Lindblom 1959; Lodge & Wegrich 2016). Additionally, Behavioural Insights tends to overlook the political dimension of scientific knowledge (e.g. Ingold

& Monaghan 2016; Weiss 1993; Cairney 2017). Many public administration studies have shown how evidence is handled, fabricated, selected, or shunned, in light of political interests that overshadow the instrumental function of science. Moreover, the particular hierarchy of evidence that Behavioural Insights promotes, with RCT-knowledge as the gold standard, is problematic. On a practical level, RCTs are only limitedly feasible as they are labour-intensive and technically complex methods, to be used in a politicized environment with scarce resources, bounded expertise, and a drive for swift decisions. On a methodological level, the claim that RCTs produce certain, fixed, universal knowledge is contestable (Whitehead et al. 2017). Rather than showing 'what works everywhere-and-every-time', RCTs provide provisional, isolated, 'closed-system' knowledge about 'what worked there-and-then' (Biesta 2007). Moreover, RCTs are only informative of the behavioural effects of a certain manipulation. They offer no understanding about why these effects happen, nor about how to actually use gained insights in unique local settings. Local knowledge, built up through practical experience over time, tends to be overlooked in Behavioural Insight discourse, because it cannot be easily explicated, codified, and disseminated as part of the body of behavioural insights (Parsons 2002). Behavioural Insights would nevertheless benefit from incorporating these non-experimental sources of evidence and horizontalizing its hierarchy of evidence (Rouw 2011).

However, as with the previous theme, the question is how universally valid the above-mentioned criticisms are in practice. Some initial observations at least partially belie them. To begin with, regarding Behavioural Insights' supposed lack of attention for macro-level behavioural influences, a key publication in the field – *MINDSPACE* – already states that 'sustainable changes in behaviour will come from the successful integration of cultural, regulatory and individual change' (Dolan et al. 2010: 13), thereby showing an awareness of the multi-level complexity of policymaking. Furthermore, common tools in the field – like the 'Gedragstoets', the 'Behaviour Change Wheel', or the 'Campaign Strategy Instrument' (BIN NL 2017b) – also pay at least some attention to macro-level influences on behaviour in their promoted analytical frameworks. Second is the concern about Behavioural Insights' narrow rational-instrumental understanding of evidence. At a basic level, my observations confirm the prevalence of such rational-instrumentalism. Dutch behaviour experts indeed tend to believe in the instrumental use of science and only scantily refer to the political function of evidence. In addition, the experimental method is celebrated widely in the field, and several behaviour experts appear to consider this method the most – if not only – valid way of gathering knowledge. But, in contrast with the critiques above, I also observed various behaviour experts who do not favour RCTs so exclusively, and draw from a wider range of sources of knowledge, including their own professional judgment and intuition.

To substantiate my argument that the claim of psychocracy is overstated, below I present two additional case studies of Dutch behavioural policy. The first case provides an example of how behaviour experts analyse policy problems from a wider, multidisciplinary angle. The second case shows that behaviour experts also draw from softer knowledge sources rather than just RCT-based knowledge, revealing the presence of a more pluralistic approach towards evidence use.

*Greening gardens*

Background characteristics
- Temporary project at central level
- Launched in 2016
- Executed by a small team of policy advisers and academics
- Goal: explore behaviourally informed ways to increase the number of green (and thus more climate resilient) gardens in urban areas

This case-study features a temporary policy project that was launched around 2016 by the Dutch ministry of Infrastructure and Water Management. In this project, a new team of behaviour experts at the ministry explicitly sought to incorporate behavioural insights in the area of climate adaptation. The specific aim of this project was to stimulate citizens to make their gardens 'greener' as a way to improve the climate resiliency in urban areas. A 'Climate Resilient Gardens' project team was established that included internal policy advisers, scientists, and advisers from governmental knowledge institutes. This team followed an approach that was based explicitly on behavioural insights: the 'DOE-MEE' approach (see BIN NL 2017b). The team started by doing a 'sanity check' (see Hommes et al. 2016) in which it investigated whether this problem was a behavioural problem in the first place. When the sanity check confirmed the potential of a behaviour change approach, the team produced an in-depth behavioural analysis of why urban citizens would (or would not) invest in greener gardens (see Rietkerk et al. 2016). This analysis identified relevant behavioural determinants based on available scientific knowledge, while at the same time explicating what was still uncertain and inconclusive about this case. This meticulous and contextual way of working challenges the idea that Behavioural Insights practices automatically frame a certain issue as a behaviour change problem and then quickly move on to designing behavioural interventions. Rather, this team first performed detailed analysis on how much there really was to win with a behaviour change approach. Then it made extensive efforts to produce a behavioural analysis, while recognizing the provisional and inconclusive nature thereof.

Furthermore, it is interesting to shed light on the nature of the analysis that was produced in this project. This analysis was based on a scheme that distinguished

between five factors: personal circumstances; abilities; motives; social environment; and decision-making processes (BIN NL 2017b). The first factor, personal circumstances, was said to relate to physical design but also to broader circumstances like the facilities, financial aspects, and law and regulations that surround people. This scheme helped to produce a fairly wide analysis. This analysis on one hand adopted the typical Behavioural Insights focus on micro-level decision-making processes. For instance, it recognized the socially contagious effects of 'grey' gardens which were increasing in numbers. Also, it addressed the fact that the impacts of a grey garden could be invisible to citizens; that grey gardens had short-term advantages over green gardens in terms of time and money investments; and that investing in green gardens might be hampered by 'choice stress' and a lack of 'mental budget' to perform garden activities. However, the analysis also addressed more systemic forces underpinning behaviour. For instance, it was found that wider social-cultural forces had a big impact on the choice for green gardens. People with, e.g. an older age, a female gender, a higher socio-economic status, a religion, and a western background, were more often found to have a green garden. Additionally, it was mentioned that it was important to look at the financial dimension behind this behaviour and the role of leisure time in today's society, as grey gardens tend to be cheaper and easier to maintain.

The observed examination of citizen behaviour from both a micro- and macro-level perspective belies the critique that Behavioural Insights necessarily investigates behaviour from a psychological perspective, solely in relationship to the physical environment. In this case, the examination was implicitly underpinned by a more complex, historical, and institutional perspective on behaviour change. Even though this project was still partially grounded in psychology, and acknowledging the fact that making a wider macro-level analysis doesn't per se result in macro-level structural interventions beyond psychocratic individual behaviour change approaches, it at least demonstrated a clear awareness of the multi-level complexities of human behaviour and of the challenges involved in redirecting it.

*Living Nudge Lab*

Background characteristics
- Temporary project at local level
- Launched in 2016
- Collaboration between schools, community centre, universities, and municipal officials
- Goal: experiment with public health-related behaviour change strategies in the real world.

In this case study, we make a shift from discussing the type of *ideas* that are prevalent in behavioural policy to the type of *methods* that prevail. Also, while the previous case plays out at the level of central government, the following is situated at the local level. It discusses a fairly atypical, small and temporary co-production between various local state and non-state actors. This collaboration – called the 'Living Lab' – was launched in November 2016 in the Dutch city of Utrecht. Its specific goal was to experiment with public health-related behaviour change strategies in real-world settings. The collaborators in the Living Lab were a community centre, two secondary schools, the municipality of Utrecht, and a research group doing interdisciplinary nudge-research (of which I was part, although I was not directly involved in the execution of the research for the Living Lab project). The research took place in 'Nieuw Welgelegen', a large community building that hosted two secondary schools. The particular area of experimentation was the shared school canteen area in this building. The Living Lab ran trials of nudges directed at pupils in two specific areas of behaviour change. To stimulate pupils to stand (versus sit) during school breaks, standing tables were put in a more central place in the canteen. To stimulate pupils to drink more water (versus soft drinks), tap water points with fruit-flavoured water were placed in a prominent place. Besides measuring the effectiveness of these specific nudges, this project also explored the use of public health nudges from an ethical perspective. Surveys and focus groups were done to learn about how school pupils thought about being nudged into healthier behaviour. A general conclusion was that pupils felt that their health was important to them, and that although they were responsible for their health choices, they viewed nudging as a potentially helpful resource (Kroese 2017).

The conclusions from the nudge-experiments measuring effectiveness were more ambiguous. While the trial on standing behaviour during lunch breaks yielded no significant behaviour change results, these results were at the same time challenged by the researchers, for instance, by noting that some students also went outside during breaks and that those who stayed inside might already have a stronger urge to sit. The trial with the fruit-flavoured water taps also resulted in mixed observations. While a slight increase in self-reported water drinking was noted, the reliability of this result was called into question. For instance, it was noted that during the time of experimentation less students had been coming to the canteen due to outdoor school events. Also, a significant group of pupils which celebrated Ramadan had not been allowed to drink during school time at all for 1 week. Moreover, during the research process, it had not been feasible to conduct self-reporting surveys with the same group of pupils consistently. The awareness of such confounding variables all made it more difficult to draw hard conclusions. The summarizing research report thus repeatedly noted the difficulties of running trials in a dynamic field setting and

interpreting the data generated from them. The report emphasized that these trials should be seen as pilots that provided lessons for future research projects that could then produce more robust results. Thus, the two clear-cut research questions that the project started with turned out to have more open-ended answers:

> In this report, we address two questions: 'is nudging effective' and 'is nudging acceptable'. It has to be noted that the research is embedded in the normal routines of the school program as much as possible, which makes the setting realistic but not ideal from a scientific perspective. The given conclusions will thus have to be viewed as lessons and ideas for the future instead of as hard evidence. (Kroese 2017: 5).

One can note several aspects about this Living Lab case that challenge the psychocracy claim that Behavioural Insights practices exclude valuable evidence from the policy process by prioritizing 'what works' knowledge. To begin with, while one part of the Living Lab indeed sought to experimentally assess the effectiveness of nudges, another part explicitly studied the ethical acceptability of nudges. This was accompanied by the introduction of specific methodological techniques (e.g. focus groups) and theoretical questions (e.g. political philosophical debates about state influence and citizen autonomy) that are allegedly less typical within behavioural policymaking. This widening focus may have been the result of the collaborative and multidisciplinary nature of the Living Lab – bringing together a range of actors representing a diversified set of interests, knowledge, and skills.

The Living Lab also seems reflective of the ultimately ambiguous nature of 'what works' knowledge. This is evidenced by the fact that the findings from both the standing table and the fruit-flavoured water tap trials were called into question by the research group and interpreted in an open-ended way. The research report repeatedly emphasized the uncertainty of the produced experimental knowledge and problematized the possibility of turning messy real-world sites into RCT-proof laboratory-like settings. The way in which the Living Lab challenged the possibilities of hard evidence-production nuances the wholly positivist image of Behavioural Insights as firm advocate of fixed, objective knowledge. At the same time, the research report stated that future experiments *would* be able to produce hard knowledge, suggesting that some of its positivist aspirations remained intact.

Another aspect that challenges the psychocracy claim relates to the effects of the Living Lab on the participating schools. One of the school principals noted during an interview that the project had led him to embrace behavioural insights in his daily work more generally. His uptake of behavioural science however seemed different than the rationalist version for which Behavioural Insights has been criticized. That

is, he adopted a less scientific and more pragmatic approach, based on his own intuitive judgment on how basic behavioural theory might help to solve operational issues. He distanced himself from the demanding experimental 'what works' approach, and the necessity of having to collect hard scientific proof for his actions. When asked how he then knew whether his behaviour change strategies actually worked, he answered:

> You don't. That's intuition. And I think that's where it stops for us. I really mean that. I think that's where it ends for us. (…) How great would it be if we'd know that the colour 'red' triggers people's motivation to take the stairs, for instance? How great would that be? But in a different way, that's also a utopia and you'd have to experimentally assess that. You'd really have to do it: the experiment. And on a small-scale that may be possible but on the large scale it's mostly still about intuition though. (…) Maybe it's the case that… you could do some literature study (…) but it remains a cost–benefit matter. (…) There are many other things that are very nice and interesting. And the benefit for me already is a change in my way-of-thinking. And that's enough.

The principal's view could be interpreted as revealing a shift from the exclusive appreciation of *episteme* to almost the exclusive appreciation of *phronesis* knowledge (Parsons 2002). As a self-proclaimed 'believer in pragmatism', the school principal practiced his own, adapted, experiential form of choice architecture: a form that was loosely informed by behavioural insights but beyond that underpinned by intuitive judgment, creative thinking, and satisficed decision-making. It is interesting to note the perceived split between academic behavioural practice and the everyday operational work. In the principal's eyes, while scientific research may help to gain specific knowledge about the effects and working mechanisms of interventions, the idea of turning the school into a scientific laboratory is perceived to be, in his own words, a 'utopia'. Scientific practice is regarded to be incompatible with his everyday work because he lacks the time and resources to go into theoretical detail or run experiments. Also, scientific study would in his eyes produce data that are less relevant at the operational level. The principal's embrace of behavioural insights shows that this initially academic approach may be locally adapted into more experiential versions, suggesting that behavioural policymaking doesn't necessarily rationalize and psychologize the policy process as deeply as some critics claim.

Overall, the Living Lab forms a small scale but nonetheless interesting alternative case of behavioural policymaking based on a collaboration between actors in government, academia, and civil society. In this collaboration, different understandings of the role of evidence became apparent that challenged the alleged psychocratic nature of Behavioural Insights. The research group, while indeed prioritizing RCT-knowledge,

nevertheless recognized the inconclusive nature of such knowledge and the methodological hardships of running experiments in real-world settings. An even more diverging understanding of the role of evidence was visible in the experiential choice architecture of the school principal, who relied on practical intuition rather than hard scientific knowledge.

## 8.5 Conclusion and discussion

Contemporary governments have initiated a behavioural turn in policymaking. This turn has had a polarized reception. Advocates point to the widespread failures of non-behaviourally informed policies in generating social change. Adversaries argue that the emerging behavioural state threatens to revive technocracy, and more particularly psychocracy, at the expense of citizen involvement and non-psychological sources of expertise. In its focus on frontstage developments in the field, the debate however overlooks a vast backwater of emerging Behavioural Insights practices. Grounded in ethnographic fieldwork over the course of 4 years, this chapter has therefore zoomed in on various lesser-known cases that have been developed in the peripheries of the field. A general conclusion is that behavioural public policy is more complex, fragmented and dynamic than the academic debate suggests. Neither the advocates nor the adversaries in this polemic debate have yet taken account of what the behavioural state currently *also* is. The advocates tend to disregard a range of critiques (concerning power, politics, and expertise) that operate underneath the current behavioural turn in policy. The adversaries tend to view the behavioural state as a coherent and finalized project, overlooking its diversity and dynamism. In addition, the radical tone of their critiques is at odds with the sometimes pastoral and incremental nature of behavioural public policy in its current form.

With regard to the claims of technocracy in general and psychocracy in particular, my conclusion would be that they are overgeneralizations. Four cases have been presented to reveal a more nuanced if not contradictory picture to some of the grand claims in the academic debate. Table 8.1 presents the specific empirical patterns reconstructed during the ethnographic fieldwork. To an important degree, it can be read as a concluding summary of the findings.

Importantly, I am not dismissing present concerns about technocracy or psychocracy altogether. The ambition of this chapter was to show that these concerns do not apply *universally*. The presented case studies demonstrate that behavioural policies are neither intrinsically technocratic nor psychocratic. Rather, it is the producers of such policies who, in interaction with the wider institutional context, make such policies more (or less) technocratic and psychocratic. At the same time, the wider

representativeness of the case studies in this chapter should not be overstated. The studied cases were purposively selected *because* they appeared to form counter cases to overgeneralizing claims. Also, the studied cases concern relatively small scale, often local and partially non-standard practices that take place in the context of Dutch governance. While Dutch behavioural practices to some extent mimic Anglo-Saxon Behavioural Insights role models, symbols and codes (Feitsma & Schillemans 2019), the Dutch context also comes with meaningful particularities that may shape local behavioural practice.

Moreover, the idea ought to be entertained that, rather than simply saying that the critical camp is too radical in its worries about the upsurge of technocracy and psychocracy, it actually may be the case that such worries have already (implicitly) impacted the development of behavioural policy at a deep level. As in a 'self-denying prophecy', the deliverers of behavioural policy may have sought to address the perceived risks and concerns voiced by the critics, leading to more moderated behavioural practices. My case studies to some extent confirm this idea. Urban Nudging, for instance, was launched as a *response* to the societally undebated ways in which behavioural science has been informing governments. And it is probable that the Living Nudge Lab research team explored the ethics of their behaviour change strategies *as the result of* an ongoing academic debate about the normative legitimacy of behavioural policies.

Whether the result of a self-denying prophecy or not, and while indeed still quite possibly odd cases in a more conservative landscape, the studied practices in this chapter at least reveal a potential of behavioural policies to withstand technocratic and psychocratic tendencies. They could be interpreted as attempts to develop a more *deliberative* type of behavioural public policy. Future research that further explores in which directions the global behavioural policy landscape is evolving, and whether a more deliberative trend is taking shape, is therefore worthwhile.

**Table 8.1: Summary of empirical patterns**

| *Theme* | *Theoretical claim* | *Empirical patterns* | *Examples from case studies* |
|---|---|---|---|
| *Technocracy* | As behavioural policies are not easily observed by citizens, they are also rarely contested, debated and/or co-produced. | I Behavioural policies can both be co-produced and be made subject of public deliberation<br>II Behavioural policies can be used as second order strategies to achieve first order participatory goals.<br>III Behavioural policies can be targeted both at citizen and non-citizen actors. | I Urban Nudging aimed to open-up the nudging debate, and empower citizens with behaviour change skills: 'Urban Nudging is a democratization of the nudge'.<br>II Landelijke Aanpak Zwerfafval (LAZ) sought to 'nudge citizens into participation', exemplified by asking a behaviour change consultancy to conduct a behavioural study called *Stimulating and maintaining participation regarding littering* (Dijksterhuis & Van Baaren 2015).<br>III LAZ also sought to activate municipal officials instead of citizens with regards to litter policy. 'How do I get the municipality to engage with this [policy]? That's the first line of behaviour change that you've got in mind.' |
| *Psychocracy* | Behavioural policies seek to change behaviour through micro-level environmental redesign, and disregard powerful behavioural influences at the macro-level. | Behavioural analyses can take into account macro-level behavioural influences. | The Climate Resilient Garden project team conducted a wide behavioural analysis of garden-related behaviour of urban citizens. This analysis included macro-level circumstances like the facilities, financial aspects, cultural norms, and laws and regulations that surround citizens. |
| | Behavioural policies are underpinned by a radical reliance on experimental knowledge, which is believed to provide fixed and universal knowledge. Other valuable methods of gathering evidence are excluded. | I Behaviour experts can be aware of the ambiguities around evidence-based policy.<br>II Behavioural policies can both be derived from experimental and non-experimental knowledge. | I The Living Lab researchers repeatedly emphasized the inconclusive nature of their experiments in a setting that was 'realistic but not ideal from a scientific perspective. The given conclusions will thus have to be viewed as lessons and ideas for the future instead of as hard evidence.'<br>II The school principal noted how the Living Lab had led him to incorporate behavioural insights, although in a more intuitive way. 'How great would it be if we would know that the colour red triggers people's motivation to take the stairs (…) But (…) that's also a utopia and you'd have to experimentally assess that. (…) And on a small-scale that may be possible but on the large scale it's mostly still about intuition though.' |

*Chapter 9*

# Interdisciplinizing the Behavioural State

*A Behavioural Insights movement has flourished within the global policy realm. While this movement has been deemed interdisciplinary, incorporating behavioural science theories and methods in a neoclassical economics-governed policy process, this chapter analyses the bounded form of interdisciplinarity that characterizes it. We claim that an engagement is missing with the broader sweep of social sciences, which share similar concerns but deploy different analytical perspectives to those of Behavioural Insights. Focusing on two central concepts (context and evidence), we aim to show how Behavioural Insights' bounded interdisciplinarity implies limited understandings of context and evidence, thereby undermining its complex problem-solving potential. At the same time, we highlight some alternative examples of behavioural practice that do explore new critical interdisciplinary horizons.*

---

Chapter 9 is co-authored by Joram Feitsma and Mark Whitehead. It is based on an earlier version presented at the annual European Consortium for Political Research conference in 2018 in Hamburg. It is currently under review for publication in an international academic journal. Joram Feitsma and Mark Whitehead have contributed about equally to this chapter across all research stages. They have collaborated extensively in designing the study and drafting the manuscript. In terms of the presented analysis, Joram Feitsma is largely responsible for the section on the role of evidence in Behavioural Insights, while Mark Whitehead has written most of the section on the role of context.

## 9.1 Introduction

In his book *How far to Nudge? Assessing Behavioural Public Policy* (2018) Peter John charts the startling rise of behaviourally informed public policies. Fusing the insights of psychological, behavioural, and economic sciences, Behavioural Public Policies are associated with the development of governmental interventions in areas of public health, environment, and personal finance (*inter alia*), which are predicated on more empirically grounded, and realistic, understandings of human motivation and action (see John 2018; Oliver 2017; Thaler & Sunstein 2008; Whitehead et al. 2017). In the opening chapter of *How far to Nudge*, John argues that the science behind Behavioural Insights (hereafter BI) is 'genuinely interdisciplinary' (2018: 4). We take the notion of *genuine interdisciplinarity* as our point of departure. John's statement is of interest to us because it simultaneously is unquestionably true, while also raising important questions about the particular form that interdisciplinarity takes in relation to BI. In the analysis that follows, we critically analyse the interdisciplinary nature of the sciences that inform behavioural public policies and suggest that it is characterized by a distinctly *bounded* mode of interdisciplinarity. This chapter considers the practical implications of the bounded interdisciplinary form of BI, and the more general theoretical implications of bounded understandings of interdisciplinarity for critical analyses of interdisciplinary endeavour.

We understand BI as a distinct community of agents and aligned practices, embodying a coherent package of symbols, ideas, methods, and political and ethical outlooks (Ball & Feitsma 2018). A crucial trait of this community is its use of techniques that are neither regulatory nor market/incentive based. Instead, BI utilizes psychological insights (largely drawn from behavioural and cognitive psychology) into the nature of human decision-making in order to develop prompts to action that are easy to resist (see Thaler & Sunstein 2008). To put things another way, BI exemplifies the application of psychological governance within free societies (Whitehead et al. 2017). Within this chapter, we are broadly supportive of BI's attempts to devise programmes that challenge the dehumanized assumptions of market rationality that have informed neoliberal government for the last thirty-five years (Davies 2014). Notwithstanding this, our primary goal is to contribute towards, and extend, critical theoretical perspectives on this movement (Feitsma, 2018a; Furedi 2011; Leggett 2014; Lepenies & Malecka, 2016; Davies 2014; Mulderrig 2017a, b; Pedwell 2017).[1] We ultimately claim that the interdisciplinary zone associated with BI

[1] Existing critical analyses have so far tended to focus on the ethical issues of related policies (with particular concerns being raised about acts of unconscious manipulation; a lack of tolerance; and demeaning assumptions about human nature), and the political implications of associated regimes of behavioural government (in particular the depoliticization of public

is characterized by a pronounced lack of engagement with a broader range of social sciences, which share similar concerns, but employ very different analytical and methodological perspectives (e.g. Jones et al. 2013; Shove 2010; Mols et al. 2015; Pykett 2018; Lepenies & Malecka 2018). There are many ways of demonstrating the critical contribution that the broader social sciences could make to behavioural policymaking. In this chapter, we focus on the specific contributions that are evident within our own disciplines of human geography and public administration. In the analysis that follows, we focus on two concepts: the geographical notion of *context*; and the public administration question of *evidence*. As we see, questions of context and evidence are central concerns, indeed leitmotifs, of BI. In our analysis, we aim to show how the *bounded interdisciplinarity* of this community can result in limited understandings of context and evidence, which arguably decreases its ability to solve complex public issues. At the same time, we also aim to show how this critique does not universally apply as in some parts of the field an encouraging *unbounding* of BI's interdisciplinarity can be observed.

The arguments presented in this paper are grounded in long term ethnographic fieldwork stemming from two independent research projects on BI. For one of the projects, over 100 qualitative interviews have been carried out with BI markers, experts, and practitioners in six countries around the world, over a nine-year period. The other project is grounded in four years of fieldwork on BI experts in Dutch central and local government, carrying out 36 interviews and conducting over 475 hours of participant-observation in the field. To empirically substantiate our analysis of BI, various case studies have been drawn from these research projects. In selecting these case studies, we followed a logic of theoretical/purposive sampling, purposefully screening for cases that helped us in showing the nature of the field's interdisciplinarity, and particularly the *varieties* that exist herein. We thus also purposefully screened for counterexamples that ran against the initial main claim of bounded interdisciplinarity. We present these counterexamples to show where and how the field is to some extent already making critical interdisciplinary excursions. Due to this purposive sampling strategy, these cases should not necessarily be taken as representative for the field in a wider sense.

## 9.2 Behavioural Insights policy in critical interdisciplinary context

### *The interdisciplinary origins of Behavioural Insights*

In his personal history of BI – *Misbehaving: The Making of Behavioural Economics* – Richard Thaler (2015) provides valuable insights into the interdisciplinary origins of

policy and the over-extension of state influence) (see Furedi 2011; White 2013; although for a sense of broader critiques see Jones et al. 2013 and Pedwell 2017).

the movement. Thaler describes how in the 1970s he began compiling a list of peculiar human behaviours: the kinds that were inconsistent with economic theory. In trying to make scientific sense of these behaviours (the kinds that see us favouring the present over the future, status quo over change, and the behaviours of others over our own judgement), Thaler describes a chance encounter that led him to become aware of the pioneering work of Daniel Kahneman and Amos Tversky. For several years, Kahneman and Tversky had been attempting to bring psychological explanations into economic theory (see Kahneman et al. 1982; Kahneman 2011). Their research suggested that humans lack the time and cognitive bandwidth necessary to consistently behave in rational ways (Camerer 2001). This body of scholarship also showed that if economists wanted an account of actually existing human behaviour it appeared they needed an interdisciplinary dialogue with certain branches of the psychological sciences. The empirical work of Kahneman, Tversky, Thaler (and others) generated a fusion between cognitive psychology and economics that would eventually become the interdisciplinary field of behavioural economics. Behavioural economists now provide the scientific bedrock for BI and dedicate their time to better understanding the *bounding* of human rationality.

It is important to recognize that behavioural economics was not the first, or only, form of interdisciplinary engagement between the psychological and economic sciences. Sent (2004: 738) has identified early, pre-Freudian, interdisciplinary dialogues between economics and psychology in the mid-nineteenth century (these interactions were based on the economic application of psychological notions of *sensation, stimulus and response*. Economists' rejection of Freudian accounts of unconscious behavioural motivations, and Milton Friedman's powerful defence of the rationality principle, did, however, stifle interdisciplinary dialogue during large parts of the twentieth century (ibid). The renewed interdisciplinary engagement between psychology and economics in the post-war period is linked to two things: 1. The cognitive revolution in the psychological sciences; and 2. The pioneering work of Herbert Simon on *bounded rationality* (Simon 1957; Sent 2004: 739; see Strauss 2009). According to Sent (2004), the cognitive revolution in psychology and the work of Herbert Simon facilitated the emergence of an *old school of behavioural economics* which drew together a series of scholars working at the interface of psychology and economics. Sent argues that this old school was characterized by a 'dismissal of the mainstream [economic] focus on profit and utility maximization and equilibrium as well as an effort to develop an alternative' (ibid.: 741). At the heart of the old behavioural economics school project was a desire to replace the figure of *homo economicus* with a new, empirically grounded, theory of human subjectivity.

The most significant and sustained form of twentieth century interdisciplinary dialogue between psychology and economics was, however, facilitated by a different group of behavioural economists. According to Sent (2004: 742), the *new behavioural economics school* emerged out of the 1970s work of the aforementioned Daniel Kahneman, Amos Tversky, Baruch Fischoff and Paul Slovic. This new school deviated from the older cohort of behavioural economists to the extent that 'they started from the rationality assumption that has characterized mainstream economics and next analysed departures from this yardstick, as opposed to developing an alternative one' (ibid: 743). It is difficult to discern precisely why new behavioural economists should have sided, broadly, with the rationality assumption. As Sent astutely observes, '[w]hile new behavioural economics seeks to strengthen mainstream economics, there is nothing inherent in appeals to psychological insights that requires this' (ibid.: 750). It could reflect a normative commitment to the notion that rational action is right, and thus, even if it doesn't reflect actually existing behaviour, it must still be assumed as a goal. Alternatively, it may speak to a more theoretical presumption that rationality can still explain much in economic life, even if not everything. What is clear, is that in taking rationality as the criterion against which psychological studies of the irrational should be measured, new behavioural economics sought to increase the explanatory power of economics and suggest that new forms of explanatory power for human conduct need not be forged too far outside of the discipline (ibid.).

The complementary alignment of certain branches of cognitive and behavioural psychology and economics forged within the new behavioural economics school was significant to the form that subsequent interdisciplinary inquiry within BI would take. The work of the old school of behavioural economics suggested that novel theories of human judgement, which sat clearly outside of existing economic thinking, were required, thus inviting potentially more open forms of interdisciplinary inquiry. The new behavioural economists developed a more instrumental form of interdisciplinarity within which psychology was used to explain and correct the specific instances where economic theories of rational human action proved to be consistently erroneous. Significantly, in the context of this chapter, this form of heavily circumscribed interdisciplinarity was based upon an increasingly experimental form of economics, which sought to use verifiable psychological methods to iteratively refine economic modes of explanation. To these ends it is clear that the interdisciplinary form of behavioural economics (and by extension BI) was about the alignment of economics with very specific branches of psychology (particularly the cognitive and behavioural traditions). From its very initiation, the scientific basis for BI was effectively excluded from significant chunks of the very discipline it was pairing with. While interdisciplinarity rarely involves the functional integration of entire disciplines, it is clear that BI is based on scientific dialogue that

has externalized certain psychological insights into the nature of human emotionality and irrationality. This externalization processes has involved the exclusion of psychoanalytical insights into the human condition, and, more recently, made it difficult for the BI to incorporate the recognition that is being given to the social psychology of emotions (Keltner et al. 2014). As we will discuss in greater detail in the second half of this paper, this circumscribed (or bounded) interdisciplinary dialogue had critical impacts on the relationship between BI and questions of context and evidence.

***Critical interdisciplinary perspectives on Behavioural Insights***

In order to develop a critical perspective on the interdisciplinary form of BI, it is necessary to consider the very nature of what it is to be interdisciplinary. In general terms, interdisciplinary activity is defined as research that involves the '*integration* of knowledge originating in two or more [academic] fields' and which is distinct from cross or multidisciplinary work, which involves '*contributions* from two or more fields to a research problem' (Jacobs & Frickel 2009: 45). The distinction between knowledge *integration* and knowledge *contributions* suggests a demarcation between two forms of academic collaboration. While integration suggests the *a priori* synthesis of research insights and techniques in order to openly explore a knowledge problem in new ways, knowledge *contributions* are typical of breaking down a research problem into discrete parts that can be explored by different disciplines while avoiding any sustained engagement between those disciplines (MacLeod & Nagatsu, 2018). According to MacLeod and Nagatsu (2018), interdisciplinarity is also characterized by three additional features: a *focus on real world problems*; the development of *more comprehensive perspectives* on problems; and a fundamentally *disruptive element* (ibid: 2-3). The focus on real world problems is a feature of interdisciplinary endeavours because by definition these are problems that do not arise within the confines of existing disciplinary purviews. The comprehensive nature of interdisciplinarity is associated with a 'more holistic or systematic understanding of a set of phenomena' (ibid.: 3), and related attempts to position a problem within a more complex web of interpretation. Interdisciplinary inquiry's *disruptive* nature is a feature of modes of analysis that lead to the destabilization of existing disciplinary modes of analysis and the formation of novel lines of academic investigation (ibid).

We acknowledge that the BI movement can be considered interdisciplinary, particularly in relation to its *focus on real world problems* that exceed established disciplinary concerns (namely the irrational behaviours of economic actors) (see Sanders et al. 2018). Crucially, however, we think that on the criteria of *comprehensiveness* and *disruptiveness* the degree of interdisciplinarity is considerably limited. In its tight embrace of certain disciplines and subsets thereof – reflecting the

particular interdisciplinary engagement between cognitive and behavioural psychology, and economics – BI overlooks a whole set of other disciplinary fields, including cognitive anthropology, cultural studies and other branches of psychology, as well as our own fields of human geography and (critical) public administration.[2] Table 9.1 outlines the interdisciplinary forms of engagements that characterize BI, and the complementary, but largely unintegrated, disciplines from which additional perspectives could be gleaned. It shows how BI is interdisciplinary in certain ways, but simultaneously points to the scientific territory in relation to which it appears to lack substantive interdisciplinary engagement (also see Lepenies & Malecka 2018; Pykett 2018; Jones et al. 2013). Notwithstanding BI's valuable contributions in refining *cognitive* and *behavioural* accounts of human decision-making, its interdisciplinarity thus seems to be of a relatively *comfortable* kind.

It is difficult to systematically evidence our claim that BI is characterized by a form of bounded interdisciplinarity (particularly given the growth of related research in this area over the last decade). We do argue, however, that there are several reasons to think that this claim may be valid. First, it is important to establish that while the theories of behavioural economics, upon which BI is based, have become increasingly heterodox, we claim more specifically that it is BI itself that has tended towards more limited interdisciplinary engagements. Writing in this journal, for example, Hansen (2018: 192) has observed that in attempting to translate the ideas of behavioural sciences into policy solutions, BI has often been guilty of 'labelling behavioural phenomena with behavioural terminology from the desktop' of behavioural economics. According to Hansen this tendency has impeded more sustained, diagnostic research into the nature of real-world problems (ibid). This is precisely the form of applied field research that would open up the study of behavioural problems to a broader spectrum of disciplinary perspectives and enable the development of new conceptual vocabularies. It is important to note, of course, that the particular interdisciplinary bounding of BI around well-established core principles of behavioural economics has as much to do with the political pressures for quick wins as any particular interdisciplinary bias. Beyond this specific claim, however, it is also clear that the interdisciplinary space defined by behavioural economics, and the broader field of economic psychology, is characterized by relatively constrained

[2] The recent emergence of the field of 'behavioural public administration' (Grimmelikhuijsen et al. 2017), merging insights from behavioural science and public administration, demonstrates however that the field of public administration is certainly not entirely excluded from BI's knowledge content. Rather, it is the more critical strands of public administration – those that problematize trends such as rationalization, scientization and psychologization – which appear to be eclipsed.

levels of interdisciplinary engagement. In a survey of published work in the Journal of Economic Psychology, for example, it was recognized that while undoubtedly interdisciplinary in scope, contributory authors were overwhelmingly from the behavioural and economic sciences, with additional inputs from the fields of consumer and market research (Kirchler & Hölzl 2006). Finally, we have presented our theory of the bounded interdisciplinary of BI to a senior figure in BIT UK who suggested that it reflected a reasonable approximation of the field.

**Table 9.1: Forms of disciplinary integration and non-integration within BI**

| | *Disciplinary Integration* | *Indicative Disciplinary Non-integration* |
|---|---|---|
| *Limited cognitive bandwidth* | Economics; Cognitive Psychology; Neuroscience | Cultural Studies (Non-Representation Theory); Cognitive Anthropology |
| *Social context of decision-making* | Economics; Social Psychology; Behavioural Psychology | Anthropology; Sociology; Political Sciences; (Critical) Public Administration; History; Cultural Studies; Social Identity Theory; Complexity Science; Innovation Studies; Science and Technology Studies |
| *Environmental/material context of behaviour* | Economics; Cognitive Design; Behavioural Psychology; Neuro-Architecture; Consumer Research | Human Geography; Anthropology ; Science and Technology Studies; Cognitive Anthropology ; Complexity Science; Innovation Studies |
| *Realities of everyday lived experience* | Cognitive Psychology; Behavioural Psychology; Economics; Social Psychology | Existentialism; Phenomenology; Anthropology |
| *Change strategies and human engineerability* | Economics; Behavioural Psychology; Cognitive Design; Social Marketing | (Critical) Public Administration; Innovation Studies; Complexity Science; Human Geography; Science and Technology Studies; Sociology; Political Sciences |

## 9.3 Behavioural Insights, public administration, and the quirks of 'what works'

Having established a broad overview of the interdisciplinary integration and non-integration between scientific disciplines within BI, the paper now zooms in on a particular non-integrated interdisciplinary territory. Specifically, it brings into view insights from (critical) public administration. A classic theme within this field is the use of knowledge, asking the question of how knowledge comes to shape policy decisions through being produced, defined, selected and authorized in certain ways (Lindblom 1959; Weiss 1979; Nutley et al. 2010). More recently, this theme has been addressed in a particular concern with 'evidence-based policy' (Newman 2016), entailing an agenda predicated on a desire to gather *epistemic* knowledge that provides rigorous, codifiable insights into 'what works' (Parsons 2002). Although BI has emerged two decades later than evidence-based policy (hereafter: EBP), both agendas are interwoven, and it could be reasoned that they have been developing 'in symbiosis' (Einfeld 2018). In accordance with EBP, BI views the relationship between science and policy from a rationalist perspective (Lodge & Wegrich 2016; Parsons 2002; Feitsma 2018a, 2019). Policymaking is viewed as a matter of technical problem-solving through rational analysis. While the field of BI appears to prefer to improve policy incrementally – in this sense deviating from a grand rationalist theory of design – it still seeks to reach those incremental improvements in a highly rationalized way, with comprehensive analyses and rigorous experimental methodologies. BI's rationalist ambitions are perhaps most visible in its advocacy of randomized controlled trials as the gold standard of evidence production, and in how it has come to re-evangelize the rational policy cycle.

Such rationalism has been widely contested in the critical EBP literature. To begin with, the idea that it is possible to start from a shared, sharply defined problem perception is already problematic (Lindblom 1959). This especially applies to 'wicked problems', which are characterized by high levels of both empirical *and* normative uncertainty, meaning that even their status as 'problems' is disputed. In such cases there is no clear-cut behaviour change goal to begin with, and top-down attempts to formulate such an ambition anyway risk evoking resistance and further amplifying uncertainty.

Also problematic are the assumptions regarding linearity and analytical comprehensiveness in policy design that seem to be implicit within EBP and also BI, at least judged by its frontstage communication (Feitsma 2018b). Such assumptions neglect that policy decisions are made in a political-administrative context of uncertainty, limited resources, and bounded rationality (Lindblom 1959). In practice, policy analyses remain inconclusive, selective, multi-sided and do not per se follow a

linear or cyclic process. Technically complex and labour intensive evaluation methods like RCTs are only feasible to a limited extent. And even when used, such rigorous methods do not produce certain and transferable knowledge, but rather provide provisional and isolated knowledge about 'what work*ed*' (Biesta 2007; italics JF) in a specific context (Deaton & Cartwright 2018). Evidence about the specific impacts of policy solutions often remains of an inconclusive nature (Esterling 2016).

Another issue is that BI appears to make remarkably little public mention of the political nature of science-policy interaction. Generally, it seems to present a selective view on the politics within academia, which is the site of many conflicts, controversies and 'science wars' between diverging paradigms (although see John 2018). Likewise, it displays limited perception of the politics within the policy arena (although see Halpern, 2015), which is affected by ideological struggles and bureaucratic 'turf-wars'. Moreover, the field calls little attention to the political dynamics occurring in the space *between* science and policy. Critical public administration scholars studying evidence use have repeatedly pointed out that policy 'facts' do not speak for themselves but become authorized as 'evidence' through filtration processes (e.g. Strassheim 2017; Cairney 2016). Policymakers – both consciously and unconsciously – fabricate, ignore, cherry-pick or dispute evidence, or filter it in other ways (Weiss 1979; Lindblom 1959; Ingold & Monaghan 2016). EBP, and so too BI, is thus best thought of as a normative ideal, or a 'useful myth' (Boswell 2018) that provides politico-epistemological legitimacy through the authority of positivist science.

The 'politics of expertise' (Fischer et al. 2015) that takes place at the science-policy-interface reflects the struggle between scientific disciplines, competing over power and legitimacy. BI's tendency towards fairly naïve objective/rationalist understandings of evidence tends to neglect this. Moreover, we claim that the bounded interdisciplinarity of BI – entailing a convenient fusing of economics with certain branches of psychology – is, in part, a product of the forms of scientific power struggle that BI's approach to evidence denies. In this power struggle, BI is continuously reproducing its own disciplinary assumptions while disempowering others (particularly the structuralist social sciences; see Jones et al. 2013). This may lead to an overpsychologized policy process, in which policy design is narrowly approached from an individualist, cognitive-economic orientation that does not thoroughly explore contexts at the structural level (Shove 2010). This risk of overpsychologization may also be reflected in a strong prioritization of RCT knowledge and the resulting disregard of a whole series of other forms of knowledge (Deaton & Cartwright 2018; Jones & Whitehead 2018). In particular, qualitative, local, tacit, and *phronetic* (Parsons 2002) types of knowledge may get excluded.

***Reimagining the use of evidence within Behavioural Insights***

Although we would contend that a large part of BI is susceptible to these EBP critiques, simultaneously it must be noted that we have observed practices that (to some extent) have been provisionally responding to these critiques. Such practices are opening up an interdisciplinary dialogue between BI and critical public administration. To further inspire such dialogue, this section gives two illustrations of such practices, offering alternative and interdisciplinary *unbounding* (yet also in some other ways still interdisciplinary bounded) ways of gathering and valuing evidence.

*Pluralist forms of Behavioural Insights in Dutch environmental policymaking*

For a first example of alternative BI practices we take a closer look at some behavioural policy initiatives in the area of Dutch environmental policy. One actor of interest is the Behavioural Insights Team for Infrastructure and Water Management (hereafter BIT IenW), which is noteworthy due to the implicit theoretical outlook that shapes its projects. This outlook stepped outside the confinements of behavioural science and also invited insights from public administration and sociology. This was reflected in BIT IenW's analytical practice, which not only included a conventional micro-level psychologist's perspective but also a macro-level sociologist's perspective (also see Feitsma 2018a). Another part of BIT IenW's move beyond strict behavioural science was how it defined its problem cases, and consequently how it went about gathering valuable evidence. For instance, one project focused on overcoming some of the negative consequences of the distribution of goods within urban areas, particularly directed at diminishing noise and air pollution, traffic congestion and unsafe traffic situations (KIM 2017). A particular behavioural challenge revolved around getting internet consumers to pick up products at a designated location (instead of having it sent to their home addresses) as that would result in reduced traffic congestion. For this case, BIT IenW assigned an economically trained scientist specialized in urban technology to conduct a behavioural analysis (KIM 2017). Interestingly, the analysis reflected influences from a range of alternative disciplines, varying from sociology to urban studies and public administration. Notably, the policy problem was defined as a 'wicked problem' that reflected a classic dilemma of a 'tragedy of the commons' (KIM 2017). Aware of the 'wicked' nature of this issue, partly stemming from its actor-related complexity, BIT IenW refrained from identifying individualized behaviour change strategies. Instead, it focused on finding higher-level solution strategies, particularly related to which actors played a key role in optimizing urban distribution and how collaboration could be facilitated. The report was underpinned by several broad theories about 'wicked problems', policy networks, and public-private partnerships. These theories didn't so much offer specific, instrumental insights about 'what works', but rather broader, more holistic

insights about the complex nature of public policy issues and how to best address that complexity. In taking these higher-level theoretical perspectives as the guiding evidence, this team moved beyond the common practice of evidence-gathering within BI, which is to search for highly specific, micro-theoretical 'bits' of evidence about the impacts of single interventions.

Besides theoretical pluralism, one can also observe methodological diversity in some of the Dutch BI practices focusing on environmental policy. A particular example of this was offered by a behaviour expert working at the Dutch executive agency Rijkswaterstaat. When prompted about his methodological practice, he acknowledged the explanatory value of RCTs while at the same time he also downplayed RCTs as being just one instrument in the behaviour expert's evaluative toolkit. Moreover, he stressed the importance of alternative methodologies. Particularly, he combined creative thinking techniques, such as 'knowledge-based brainstorms', with the strategy of harnessing local knowledge within communities in order to develop innovative and workable solutions:

> How you make that translation [from analysis to local solutions]? That's quite a tricky one. I once organized a benchmark-session, just showing a movie of Cialdini and his principles [of social influence] and then letting the audience think about how they could apply this in [the policy area of] waste separation. You'll notice that they get through the movie quite easily and that applying each behavioural change technique to the specific situation and the identified "customer journeys" yields some nice results. So that's one nice way to do it. Those people also know the target group so they're well able to think along. [It is important to] engage them with these techniques and let them think along about how to modify those or assess whether they work or not. That they develop a sensitivity for it. I don't think you can do all of that by yourself.

The practices of BIT IenW and the expert at Rijkswaterstaat reflect a degree of pluralism regarding the use of theories and methodologies. In how evidence was gathered and valued in these cases, one can see a move beyond the typical BI tendency to only search for micro-theoretical behavioural evidence about 'what works'.

*BIT UK's 'Behavioural Government' report: a convincing pre-mortem of BI?*

For a second example, we consider a report, titled 'Behavioural Government' (Hallsworth et al. 2018), which has recently been published by BIT UK. Interestingly, this report breaches with frontstage conventions by explicitly engaging in debate with

the critical EBP literature. More specifically, the report takes seriously the critique that has been aptly formulated as the 'rationality paradox' (Lodge & Wegrich 2016), holding that while the field is all too aware of the bounded rationality of the *citizen*, it shows peculiarly little recognition of the bounded rationality of the *public official*. 'Behavioural Government' therefore *does* begin to look into government's own bounded rationality. To do so, it reviews the various sorts of biases that have been found to mark the policy process. It arrives at a whole list of biases that are structured along the policy stages of *noticing*, *deliberating* and *executing*. The review draws on a wide range of academic, mostly experimental studies, in the area of 'behavioural public administration'. After identifying these biases, the report reflects on ways to overcome them to the extent that is possible. For instance, to overcome the confirmation bias of officials to tend to only take in evidence that supports the views that they already hold, one advised counterstrategy is to '[b]uild in opportunities to change course and revisit assumptions' (Hallsworth et al. 2018: 12). To prevent group-think, the team advises us to '[a]ssemble teams that are cognitively diverse' (ibid.). Furthermore, a case is made to '[b]uild trials and variations into policy execution wherever possible' (ibid.) which can generate strong evidence to counteract unrealistic assessments of instruments' effectiveness.

'Behavioural Government' is innovative in how it shifts the field's common locus of analysis. Instead of analysing the citizens and businesses that are subjected to government policies, the government itself becomes the locus of analysis. Through this shift in focus, this variant of BI shows itself to be more reflective about the ways in which decisions made by government officials are shaped by their personal background, wider socio-material context, and varied unconscious choice processes. In exploring these issues, the report engages in a broader interdisciplinary dialogue. This dialogue is also embedded institutionally, as 'Behavioural Government' has been written in direct collaboration with critical public administration scholars, including Paul Cairney, Martin Lodge and Kai Wegrich. 'Behavioural Government' might even be furthering the critical EBP agenda in ways that this program itself has thus far not been able to. That is because the report has managed to translate a diffused set of academic theories about bounded rationality within government into a more unified, intelligible, and policy-relevant model, persuasively stacked with a rich body of experimental evidence. The report is, in this sense, 'brokering' EBP critiques, moving them into the policy sphere, and translating them into the policy lexicon.

Nonetheless, there are some other aspects in which 'Behavioural Government' seems more limited in terms of its critical interdisciplinary engagement regarding evidence use. A first limiting point is that the report shows little acknowledgement of the political drivers that impact how evidence is taken up in the policy process. The report

makes an explicit distinction between 'programmatic' and 'political' policy evaluation. Programmatic evaluation means 'looking at observable costs and benefits to society, and comparing the policymaker's original intention with the eventual outcome', while political evaluation is about 'the way policies are being perceived and debated among their stakeholders' (Hallsworth et al. 2018: 16). The report then informs the reader that it has chosen to adopt the programmatic perspective. From a critical point of view, however, this programmatic/political distinction is in and of itself problematic. It underestimates how deeply intertwined politics and policymaking are (Strassheim 2017). Seemingly neutral programmatic evaluation of 'observable costs and benefits to society' inevitably has to be selective in practice in terms of what costs and benefits it researches, for whom, and how it does so. Forced to be selective, all sorts of choices must be made, all situated within a political context of value conflicts between actors, interests and ideas. Behind constructs like 'observable' and 'society' lie political questions about *who* gets to decide what these constructs mean and what they in- and exclude. As even such basic elements are already value-laden and evoke power struggles: programmatic evaluation cannot escape the territory of 'the political'.[3]

Moreover, the critical interdisciplinary dialogue could have been deepened if 'Behavioural Government' had moved beyond conventional theoretical and methodological outlooks. In terms of methods, the report heavily relies on experimental studies, implicitly conveying that non-experimental types of studies yield no valuable evidence on this. The opposite is actually true, as, for instance, cross-sectional surveys (admittedly, the report includes one small survey) and in-depth qualitative studies (e.g. McGoey 2012; Nutley et al. 2010; Boswell 2017) can offer insights into governmental evidence use in ways that experiments cannot. In terms of theories, 'Behavioural Government' also remains somewhat tied to its interdisciplinary comfort zone, as it is researching governmental biases by and large from a psychological-economic, individualist perspective, zooming in on the micro-level choice processes of individual agents, rather than exploring limits to rationality in the context of wider governmental structures and broader societal developments at the macro-level (see Sanders et al. 2018, for an acknowledgement of the social limitations of BI). Admittedly, 'Behavioural Government' (2018: 13) does acknowledge the role of wider structures at several points, stating that '[r]eforms

[3] The report's problematic attempt to depoliticize policy evaluation represents a departure from the earlier 'Policymaking in The Real World' (Hallsworth et al. 2011) report by the Institute for Government, which was co-authored by two authors of 'Behavioural Government'. In the earlier report, it *was* fully acknowledged that 'policy making can never be extricated from politics' and that 'evidence and analysis is never "pure" or above politics' (Hallsworth et al. 2011: 82).

cannot focus on individuals in isolation – they also need to consider how systems, processes and institutions create behaviours'. However, this line of thought has not informed the report in a deep sense.

To conclude this case-study, and to stick with the report's own terminology, 'Behavioural Government' could be seen as a first attempt in conducting a 'pre-mortem' of BI. The report has explored how critiques of BI's rationalist assumptions might imply a collective failure of the field and how such failure could be avoided. This deliberative turn within BI has to be applauded. Nevertheless, we have also pointed out aspects in which the report's interdisciplinary inquiry remains unconfrontational.

### 9.4 Behavioural Insights, geography, and the issue of context

Our second set of reflections on the bounded interdisciplinarity of BI concerns the issue of context. Context is a particularly important concept within BI. First, the decision-making context is a crucial lens through which a BI perspective seeks to observe and understand bounded rationality. While bounded rationality is, in part, seen as a product of the internal limitation in the cognitive bandwidths of individuals, behavioural contexts (whether it be canteens, streets, or supermarkets) are also seen to limit optimal decision-making. This emphasis on the role of context represents a crucial departure from neoclassical economics (and, indeed, liberal political theory), where economic decision-making appears to operate in a context-less non-space of desert-island-like solitude. Second, context is important within BI to the extent that it offers a key vector in and through which behavioural polies are themselves pursued (see Carter 2015; Jones et al. 2013; Whitehead et al. 2018). BI is consequently predicated upon how shifts in social, material, and increasingly digital contexts (including places, buildings and data platforms) of decision-making can offer the basis for effective shifts in recalcitrant habits and behaviours. It commonly targets *choice architectures* and environments as a basis for behavioural government.[4]

[4] Notably, the contextual dimensions of BI are actually the product of a fairly informal interdisciplinary dialogue between the behavioural and design sciences that flourished from the 1970s onwards. While early behavioural economists were primarily interesting in the internal cognitive bases for irrational decision-making, the putative fields of cognitive design and engineering actively considered the environmental bases of behavioural biases (Norman 2002; Whitehead et al. 2017: 66-67). BI would ultimately become the interdisciplinary amalgam of certain psychological and economic insights into human cognition with a dose of the design sciences' understandings of the contextual world. Given the design sciences' concerns with

The purpose of this section is to consider the potential practical and intellectual implications of extending the interdisciplinary horizons of BI in relation to issues of context. To these ends we focus specifically on existing work within the geographical sciences. Work within human geography has a long association with questions of context. Indeed, the very nature of the geographical discipline emphasizes the importance of developing spatial and temporal perspectives on various social, economic and political issues. Given the significance of context within BI, it is notable that the geographical sciences sit outside the interdisciplinary interactions associated with BI (Whitehead et al. 2018). However, it is important to note that geographical critiques of the contextual assumptions of BI have already emerged. At a general level, Strauss (2008: 143), argues that within behavioural economics the notion of context '[…] is underdetermined and remains largely untheorized'. In his analysis of public health initiatives that deploy BI, Carter (2015) has argued that while related policies focus on important local social and infrastructural contexts, they tend to ignore broader contextual processes related to class and race. Hence, it is in relation to the at best simplified, and at worst underdetermined, application of contextuality within BI that we claim geography could productively contribute to this interdisciplinary field. According to Strauss (2009: 308-309),

> [a] geographical conception of context as the decision-making environment encapsulates the permeable and mutable scales implicated in the decision-making 'moment'. Thus, the articulation of space and place as part of the conceptual working through of the notion of context must include the scalar range of individual experience: from the individual to the global, from the intimate to the distanced, from embodied to disembodied forms of experience.

From a geographical perspective, the contextualization of phenomena (whether it be human habits or neighbourhood redevelopment) is about more than merely situating them within local processes. Within geography, context is routinely understood as the particular coming together of local and supra-local forces to shape human behaviour and social processes in space (see Whitehead et al. 2018). This approach to context doesn't mean that an appreciation of the role of psychological processes and design environments offered within BI is not important. Rather, it suggests that the contextual drivers of human decision-making cannot be determined, or hoped to be changed in the long term, through a focus on local design environments alone. Critically, within geography, a contextual perspective is used to understand the spatial differentiation of things like human behaviour. Within BI, the contextual

proximate design objects and highly localized contexts, this interdisciplinary dialogue was again predicated on the forms of bounded interdisciplinarity that characterize BI.

project is currently more about understanding the generic impact that context makes on human behaviour across spaces.

***Rethinking Behavioural Insights' contextual project***

We now consider the implications of a geographical perspective on questions of context to the study of behavioural phenomena. In doing so we also consider existing attempts within BI to broaden its contextual horizons, and how these are nonetheless limited by the bounded interdisciplinarity of the project.

*Acknowledging the non-immediate context in organ donor registration policies*

The contextual limitations of BI are already evident in one of its most celebrated policy initiatives. Opt-out organ donation registers have become something of an emblem for behavioural policymaking (see Thaler & Sunstein 2008). They are predicated on the notion that while the majority of people express a preference for donating tissue and organs after their death, many never actually get around to opting in to organ donor registers. BI's solution to this problem has been to deploy one of its most powerful techniques – the resetting of defaults. By presuming consent for organ donation, while making it easy for people to opt-out of being on donation registers, millions of people were added to organ donor registers in nations including Belgium, Spain and Wales. But despite the putative success of opt-out organ donor registers, recent evidence suggests that they may not, at least in all cases, actually increase the number of organ transplants.

In December 2015, Wales moved to an opt-out organ donor system. Recent figures reveal that despite placing tens of thousands of extra people on the organ donor register, the policy change has made negligible difference to actual transplant levels (BBC News 4 December 2017). It is our contention that the limitations of this policy initiative directly reflect the contextual shortcomings of BI. The contextual failings of opt-out organ donation revolve around the fact that it tends to reduce donation to an individual act (or inaction) associated with being put on an organ donation register. While opt-out systems thus recognize the immediate contextual factors that shape organ donation decisions (particularly inertia and fear), they fail to account for the ways in which the 'decision-making environment encapsulates the permeable and mutable scales implicated in the decision-making "moment"' (Strauss 2009: 308). To put things another way, opt-out organ donation registers generally fail to account for the connections between joining organ donor registers and the contextual factors that lead to organ transplants. The first contextual factor to consider here is the tendency of families, in the aftermath of the loss of a loved one, to override their (presumed decision) to donate their organs. Recent research has revealed that opt-out systems will lead to more families deciding to go against the donation wishes of their kin

precisely because in such systems consent becomes less clear (Lin et al. 2018). In Spain, where more emphasis has been placed on understanding organ donation in the context of grieving families, the opt-out system has been much more successful. The second contextual factor relates to having greater support for logistically capturing, storing, and transporting organs. Without appropriate funding to support expanded organ donation systems, it is unlikely that opt-out system will be able to deliver significant changes in transplant rates.

Notwithstanding these issues, it is evident that BIs in the area of organ donation are adapting to address their contextual limitations. There is a tendency to now prefer the use of mandated or prompted choice as opposed to opt-out systems to promote registration on organ donation systems (whereby people are required to opt-in or out of registers when they renew their driving license, for example). Such initiatives are, however, still characterized by a focus on decision-making moments, which fails to account for the varied, interconnected socio-economic and emotional contexts within which human behaviour is situated. It is our contention that the narrow deployment of context within BI is a product of the tendency of both cognitive psychology and neo-classical economics to focus on the individual as their primary unit of analysis. The constricted understandings of context can thus be interpreted to be a direct consequence of the bounded interdisciplinarity of BI.

*BBP, context and the Flint water contamination problem*

An additional, salutatory, example of contextuality within BI relates to the pioneering work of the US Government's Social and Behavioural Sciences Team (hereafter SBST). SBST was created during the Obama Administration in order to support the development of BI across relevant public policy areas in the US. Significantly, the SBST worked to broaden the contextual remit of behavioural policies. The clearest evidence of SBST's attempt to bring novel contextual perspectives into BI can be seen in its work in Flint Michigan. Following a change in water source supplying the town of Flint in 2014, the town's water started to suffer from contamination. The water contamination was the product of various carcinogens and bacteria being present in the new water supply, and the lead that was entering the water following its interaction with old, and poorly maintained pipes (see Stillman 23 January 2017). The problems of water supply in Flint led President Obama to declare a state of emergency in the city in 2016. Further public health problems emerged as knowledge about the contamination of tap water spread. As people became wary of using tap water to wash their hands, there were outbreaks of shigellosis (a highly contagious bacterial infection) (see Nimishakavi 10 October 2016). The primary behavioural problems related to the Flint water crisis pertained to getting local communities to continue to use bottled water, to fit and regularly replace water filters within their homes, and to

maintain good handwashing habits (ibid). These behavioural goals were, however, undermined by a prevailing sense of social injustice surrounding the crisis (with certain communities having had their lead pipes replaced more quickly than others), and a general sense of mistrust towards government guidance concerning appropriate forms of water behaviour (government advice had initially suggested that Flint's water was safe to drink) (Stillman 23 January 2017).

In relation to the social justice and governmental mistrust issues surrounding the Flint water crisis, the SBST advocated a new approach to behavioural context. This new approach immediately sought to move beyond a concern with the individual psychological and economic drivers of human decision-making, to understand the role that community norms and cognition play in shaping individual decision-making. The SBST sought to work with various community groups to better understand why residents might be unwilling to accept and act upon advice they received from governmental officials regarding water filter maintenance and the safety of using tap water for handwashing. What is interesting in relation to the work of the SBST in Flint, is that in trying to understand the contextual challenges of behaviour change in circumstances of systemic social injustice and governmental mistrust, new approaches to evidence gathering were deployed. The SBST worked closely with community groups to develop detailed, qualitative insights into their perceptions of water use and the water crisis (Stillman 23 January 2017). There was a clear sense in this situation that the gathering of trial-based, generalizable, and quantitative data was neither feasible nor intellectually beneficial. It was clear that in trying to understand the contextual production of social injustice and suspicion at various local, state and federal levels, a more open-ended, ethnographic methodology needed to be deployed (context is thus not only important in terms of epistemological explanation, but also in terms of more mundane issues of methodological practice).

The work of the SBST in Flint is emblematic of a BI initiative that is attempting to develop a more multifaceted understanding of the role of context in shaping human behaviour. This attempt reflects emerging work within human geography on the local production of systems of water cognition (Wolfe 2012), as well as work within anthropology on the creation of systems of local meaning and practice. To these ends, the SBST's research embodies a real-world inspired unbounding of the forms of interdisciplinary collaboration that have characterized BI. This unbounding is, however, only partial. Although the SBST have sought to understand water behaviours in Flint through more in-depth study of how variously scaled economic, governmental and infrastructural processes have shaped community cognition, these insights have not been used to promote the development of more radically-oriented contextual strategies for behaviour change. The SBST recommendations for

behavioural policy in Flint ultimately saw a shift from multi-scalar contextual explanations of 'irrational' responses to water use advice in Flint, to the more familiar use of local social and material contextual tools to promote desired behaviour. Thus, the SBST promoted the use of fairly routine 'implementation prompts' and 'commitment devices' for water filter use and replacement, and 'social norming'/peer pressure, and greater policy transparency (to overcome mistrust) (Stillman 23 January 2017). While these may be sensible contextual strategies in some ways, they tend to result in context moving from a multidimension horizon for understanding human behaviour, to more localized strategies for shifting behaviour. So ultimately, the SBST's work in Flint initially offered a more unbounded set of interdisciplinary perspectives on the contextual production of behavioural irrationalities, only to see context be wheeled out as a set of localized solutions to these irrationalities. The refocusing on local context in Flint is likely to offer only a short-term fix to problems of water behaviour (perhaps, no bad thing in the context of a public health crisis). More worrying is the fact that in seeking to correct irrational water behaviours, such work could undermine the political energy (and logic) that is encoded in these behaviours.

## 9.5 Conclusion and discussion

In this chapter we have scrutinized the BI movement in relation to its disciplinary orientations. We have argued that, despite widely shared celebrations of its 'genuine' interdisciplinarity (John 2018), this movement is actually often based on a bounded and comfortable form of interdisciplinarity. In its boundedness, BI is severely limited in its ability to take account of the holistic nature of contemporary policy issues. While it might make an innovative contribution to policymaking it is unlikely to become a stand-alone problem solver. Hence, we have explored how the field could learn from insights from the more distant disciplinary perspectives of public administration and human geography, and where it is already making such interdisciplinary excursions.

However, it is our contention that beyond fusions with human geography and public administration, there is a much wider interdisciplinary domain to explore for BI. We envisage more *disruptive* interdisciplinary spaces in which fundamentally differing knowledge paradigms are confronted with one another (see again Table 9.1). We would, for instance, see it as valuable if BI would combine its micro-level psychologist's gaze with the macro-level perspective of the sociologist. Attempting to merge these gazes and their underlying paradigmatic views would undoubtedly raise conflict. One might even question whether such a merging is even truly possible, without abandoning either paradigm. Shove (2010: 1279) for instance, firmly believes that such attempted interdisciplinary projects are 'doomed to failure', reflecting a

methodological naivety concerning the incommensurable differences between, in her specific example, social practice theory and the behaviour change agenda. These approaches do not merely differ in their analytical gaze and instrumental kit, but already diverge in how they define policy problems (ibid.). We are, to some extent, sympathetic to this argument. Indeed a puzzling question arises about where BI might begin and end. How far could BI go in a process of critical interdisciplinary engagement and still remain BI? While acknowledging these 'boundary issues' (also see Chapter 6), we nonetheless would encourage attempts at exploring new interdisciplinary horizons and at least trying to develop critical interdisciplinarity. Even if a more disruptive interdisciplinarity would be hard to realize and also blur and destabilize BI's essence, efforts to still live up to the ideal of critical interdisciplinarity as much as possible would seem to give the field the best chances at overcoming its own epistemological limitations.

Another legitimate question might be why it would be desirable to hold onto the label and paradigm of BI in the first place, given its epistemological limitations. Why not move to a new paradigm altogether? We recognise at least two reasons to enrich and transform BI from within, rather than trying to move beyond it. First, BI still seems to be in a momentum of growth, as still more resources and support are being mobilized and new programs and units are being launched (John 2018). From an incrementalist point of view, it would seem more productive to go along with the current energy behind this institutional development and try to challenge and change it as it evolves further. Calling for a radically new post-BI paradigm from an ivory tower position would seem less fruitful, not in the least because that would also prevent policymakers from capitalizing on the unique epistemological benefits of BI. Second, as BI is still in the process of institutionalizing, now is a fertile moment to feed the field with new ideas from distant disciplinary perspectives. As the field has not yet wholly settled its professional boundaries, ample space exists for critical interdisciplinary inquiry.

Obviously, our attempt at outlining a critical interdisciplinary agenda for BI in no way means that the field will actually engage with it. Rather, various institutional, political and cognitive reasons exist for the field sticking to its current body of interdisciplinary knowledge. A widely accepted way of interpreting emerging patterns of interdisciplinary engagement, and non-engagement, and bounding, is provided by institutional perspectives (Jacobs & Frickel, 2009). From an institutional perspective, the nature of interdisciplinarity is understood to be facilitated and constrained by the organizational form of existing academic disciplines, labour markets, funding regimes, and established channels of communication (MacLeod 2018; MacLeod & Natashu 2018). Political reasons also partly explain the bounded interdisciplinarity of

the field. As part of the current neoliberal zeitgeist, and in response to no longer maintainable welfare state systems, governments have constructed a new social state-citizen contract in which the responsibilities for social problems are increasingly shifted towards citizens (Ossewaarde 2007). Strong importance is attached to the notions of individual agency and choice. More bounded interdisciplinary strains of BI are in this light politically appealing because they closely match this neoliberal narrative in terms of their methodologically individualist underpinnings (Shove 2010). It is also interesting to draw on the idea of 'strategic ignorance' here (McGoey 2012). This notion refers to the usefulness of 'knowing what not to know' in the context of policy, as 'not-knowing' may allow actors to sidestep undesired accountabilities and responsibilities. In the case of BI, we might observe a 'strategic ignorance' too. By not coming to 'know' the context at a structural level, the field legitimizes itself in not having to identify and intervene on such deeper-lying institutional, political, economic and socio-cultural factors affecting policy issues. BI's bounded interdisciplinarity lies at the bottom of this 'strategic ignorance', acting as a legitimizing constrainer, helpfully narrowing down the fields' disciplinary bounds and with that also its bounds of accountability and scope of intervention.[5]

BI's particular brand of (bounded) interdisciplinarity is not just a product of institutional and political dynamics however. MacLeod (2018: 698) asserts that *cognitive obstacles* and the 'domain specificity of expertise and scientific practice' are what ultimately determines the success, failure, and form of interdisciplinary endeavours. It is, perhaps, ironic that an appreciation of cognitive limitations should be so important to understand the confines of a branch of interdisciplinary inquiry that is predicated on the study of cognitive limitations. But it is our contention, following MacLeod, that the bounded interdisciplinarity of BI is to an important degree a product of the cognitive constraints that exist to effective forms of scientific collaboration. Academic forms of expertise are, in part, based upon systems of knowledge that are *domain specific*. Domain specificity doesn't necessarily correlate with disciplinary patterns of knowledge production, but with forms of expertise that display:

> (i) […] narrow subject matter or classes of problems [that] cognitive systems address, and (ii) [an] inflexibility given the fine-tuned

[5] McGoey (2012) asserts that strategic ignorance works best when it seems 'genuine', i.e. when the reasons for 'not-knowing' are wholly legitimate. Such genuineness can hardly apply anymore in the case of BI, as related critiques have been circulating widely in both the public and policy realm. To some extent, these have also been addressed by the field itself, particularly through 'academic broker' figures like Michael Hallsworth (e.g. Hallsworth et al. 2018).

> dedication and specialization of these systems to handling well that subject-matter alone (MacLeod 2018: 703).

Domain specificity is thus characterized by the emergence of cognitive shortcuts, which enable specialized knowledge problems to be tackled efficiently (if opaquely). The *fine-tuned specialization* associated with domain specificity in turn reflects the simplification, or bounding, of complexity so that problems do not exceed the cognitive grasps of expertise (ibid). Interestingly, theories of domain specificity are based upon similar assumptions regarding the cognitive forms of human judgements as those associated with BI. In order to function efficiently in the face of complex situations, decision-makers (both expert and lay) must take shortcuts and form habitual heuristics, which enable them to make generally effective (if sub-optimal) judgements on a day-to-day basis.

While domain specificity is often utilized to explain why interdisciplinary projects fail, it can also explain why others succeed. It is our contention that the constrained forms of interdisciplinary integration that characterise BI can be partly explained on the basis of the 'bounded habitats of knowledge practice' (Knorr-Cetina 2007: 1) associated with domain specificity. To these ends, the coming together of cognitive psychology and economics within BI doesn't reflect a form of inevitable interdisciplinary fusion, in and through which the epistemological shortcoming of neo-classical economics are resolved in the only ways they could. The BI movement embodies the emergence of new perspectives on the *problem spaces of economics* (specifically in relation to limited cognition, social influence, and environmental influence), which are consistent with the *domain specificity of economics*. A more critical form of interdisciplinarity can avoid the limiting effects of scientific instrumentality through a focus 'on the recognition of limits, defining how much or how little information is needed to address a challenge at hand?' (Frodemann & Mitchum 2007: 508). More critical interdisciplinarity thus seeks to endure beyond the initial interactions of interdisciplinary dialogue, to consider not just if disciplines can be usefully combined, but also to contemplate what other, potentially more challenging, forms of knowledge and collaboration are needed to address the real-world problems towards which scientific endeavours are directed. It is in these contexts that interdisciplinarity can embrace more *comprehensive* and *disruptive* perspectives.

The BI movement continues to gain traction in the global policy sphere. Meanwhile, 'wicked problems' persevere too, begging for more comprehensive interdisciplinary approaches. Our chapter has sought to connect these two developments. Let's see whether there is a future for BI that is genuinely, and critically, interdisciplinary.

# III | Discussion

*Chapter 10*

# Futuring the Behavioural State

## 10.1 Introduction

The nucleus of an international Behavioural Insights (BI) community has surfaced, aiming to infuse the policy process with recent insights from the new school of behavioural economics. This community – and its associated set of ideas, technologies and values – has been met with markedly different responses. Advocates have portrayed BI as a great promise for benign and effective government. Some critics have depicted it as ineffectual and methodologically limited; others have attacked it for its deemed technocratic intrusiveness into citizens' lives and choices. Meanwhile, sceptics argue that this so-called 'behavioural turn' is neither innovative nor of any substantial impact.

This study has shed much-needed empirical light on this debate by mapping and interpreting the actual size, shape and impact of BI-style initiatives and units within the contemporary policy world. Instead of looking at BI's programs and policies, it has shifted the research focus to its early adopters in Dutch government and to what they are actually doing, guided by the broad research question:

> *What can we learn about the behavioural state by studying its members and their practices from up close?*

In this discussion chapter, it is time to take stock of the findings presented in the empirical chapters and – in particular – reflect on how these inform our thinking about the future development of what at the outset of this study I have coined 'the behavioural state'. This chapter addresses two trios of questions.

The first trio to be considered here recaptures the empirical questions formulated at the outset of this study (see Table 1.1). The first question – *Is there a behavioural state?* – asks about the presence, scope and novelty of the behavioural state. The second – *How is the behavioural state developing?* – asks about how this state is institutionalizing and to what extent it is forging a coherent professional program and identity. The third question – *How does the behavioural state see?* – asks about the nature of its prism on the world as represented by the kind of expertise it mobilizes and deploys.

The second trio of questions moves from stock-taking to exploring feasible and desirable futures. The fourth question – *What may become of the behavioural state?* – asks about the future that lies ahead of the behavioural state. In response to this question I explore three plausible scenarios for this state's future development. The fifth – *Should there be a behavioural state?* – shifts the ground from empirical to normative discourse. With the three future scenarios in mind, it asks how desirable the emerging

behavioural state is, and the extent to which its development could be 'tamed' so as to ensure society reaps its potential benefits while being protected from its potential risks, drawbacks and 'dark side'. Finally, the last question – *How should the behavioural state be studied?* – asks for a strategy and an agenda for future research on the behavioural state.

**10.2 Is there a behavioural state?**

Let's start with the most basic, foundational question: is there really such a thing as a 'behavioural state' today? What forms has it taken, and how established and consequential is it really? The academic debate on these matters has so far shown little consensus. BI advocates have emphasized the field's rapid development and early achievements in public policymaking. Critics have also observed the rise of an impactful (yet for some also methodologically bounded) behavioural state, which they claim might actually be/become *too* dominant. Sceptics on the other side have pointed at the field's marginal, flimsy and fuzzy nature. Viewing BI as overlapping with so many already existing ideas and practices, they claim that it doesn't possess sufficient distinctiveness and momentum to have been truly institutionalized within governments.

My surveying of the BI territory in academic research and in Dutch government suggests that indeed a noticeable and accelerating behavioural turn has occurred during the past two decades. In academia, the tenth anniversary of BI's breakthrough publication *Nudge* was emphatically celebrated (Behavioral Scientist 4 September 2018) and heralded as a robust trend instead of an ephemeral hype. In the institutional domain, as Chapters 3 to 6 have shown, numerous concrete and structural BI initiatives have been set in motion. By now there are over 200 institutions applying behavioural science according to the OECD (2017) and many specialized behavioural units, networks and expert practices have been launched across the global policy realm. These active behavioural units are continuing to forge new connections, spread and promote knowledge, develop new tools and methods, and develop the research agenda of the field. They also continue to tackle new policy issues with BI-associated methods and intervention strategies.

While the Anglo-Saxon policy realm forms the field's main foothold, BI has begun to pop up in other places across the globe (Whitehead et al. 2017). This study has shown how the BI agenda has also gained traction in the Netherlands. It has shown that the Dutch interest in BI has spread wide, with behaviour experts operating in a wide range of governmental institutions, including ministries, executive agencies, regulatory agencies and municipalities. Furthermore, Dutch experts are active in a

diversity of policy areas, including mobility, public health, food waste, traffic safety, littering, and personal finance.

On the basis of the steady and continuous growth of behavioural policy practices worldwide and my fieldwork in the Netherlands I conclude that there is no reason to expect stagnation of the BI movement in the near future. Further expansion seems more likely instead. BIT UK for instance has only been expanding in the last decade, and currently has six offices worldwide. Signs like the introduction of a formal Behavioural Insights Network Netherlands (BIN NL) at the interdepartmental level indicate that the BI field will mature in the Netherlands as well. In 2018, BIN NL organized its second annual conference ('De Dag van het Gedrag') with as its theme 'Next Level', expressing an ambition to complexify and innovate the BI agenda. The recent launch of the Behavioural Insights Group Rotterdam (BIG'R), the biggest Dutch BIT so far with (at the time of research) about 25 staff members, also is a sign of expansion rather than of decline. And as a last example, the Dutch Healthcare Authority (NZa) now employs '*senior* behaviour experts' (italics JF). A simple 'whim-of-fashion-explanation' of the BI trend would fail to acknowledge these signs of expansion.

But equally the ascendency of the Dutch behavioural state should not be overstated, as it is still very much at an early and explorative stage of development. The field by and large consists of fledgling practices that yet have to formalize, mature, and gain institutional authority and influence. This is illustrated by the fact that many of the behaviour experts I studied operated solitarily, and that all but one Dutch BI units remain relatively small - with at most five staff members. BI's 'new kid in town' status is also reflected in its outputs and achieved results. There are a number of, mostly small-scale experimental, BI projects with actual results in terms of improved policy outcomes. A report by BIN NL (2018) provides an overview of projects in Dutch central government. The report showcases 14 outcome-related results, but also lists about another fifty BI projects many of which have not (yet) produced tangible outcomes. While this list is not exhaustive, it does suggest that the field's portfolio in terms of achieved results is still modest. It is also important to note that the gains the field claims for itself so far are more concerned with throughputs (i.e. improved internal work processes) rather than with outcomes (i.e. delivered social change). The gains are to be found in the use of improved policy analyses (typically new behavioural economics informed analyses), solution strategies (often nudges) and evaluation methods (preferably field experiments).

Thus, although BI advocates may still turn out to be proven right when they claim that the behavioural state is 'here to stay', the current state of the art in the

Netherlands at least is that its alleged potential for radically increasing policy effectiveness and realizing widespread social change remains to be delivered. The BI community has not (yet) become a mature successor to the New Public Management policy paradigm. But neither has the critics' dystopian image of an overbearing behavioural state materialized. In the grand scheme of things, behavioural expertise is presently a *modest presence* within the Dutch policy establishment. It is certainly not – yet – part of its default repertoire and style.

As they are still in the process of 'setting up shop', BI units and experts don't just occupy themselves with the substantive aspects of behavioural policymaking. Rather, a fairly large part of their activities relates to entrepreneurship and institution-building, ensuring that innovative ideas and methods land within their organizations. They spend an important part of their time and energy in becoming seen as legitimate actors whose added value derives from pushing the adoption of new behavioural insights within government. Core activities are forging networks, spreading knowledge, organizing resources, and building support structures. This is why, as Chapter 3 has argued, it makes more sense to view behaviour experts as 'knowledge brokers'. They seek to ensure BI's intellectual legacy gets adopted by and translated into government institutions rather than that they directly apply behavioural expertise in concrete instances of policy design.

Just as the question whether BI experts will become fully established awaits a definitive answer, so too the question remains whether they will develop a distinctive competence (Selznick 1957). The field has certainly pushed the use of a particular body of knowledge, i.e. the new behavioural economics (NBE), which in its purest form were largely unknown in governments before the 2000s. As some BI sceptics have noticed however (e.g. Fitzpatrick 2011; Bonell et al. 2011b), there has been a long lineage of governmental practices that (at least implicitly) were 'behavioural' without drawing on NBE (See Chapter 2 for a brief history of this institutional lineage). What this wider history exactly entails, and which different waves of behavioural public policy have emerged over time, is a historical question beyond the scope of the present study, which focuses on the early adopters of the current wave. Nonetheless, my observations can shed light on why the potentially vast 'pre-history of nudge' (Vallgarda 2012) is fairly little acknowledged in the contemporary field. One thought is that BI's 'collective amnesia' about its intellectual and institutional roots makes sense if we consider that they are still in the process of institutionalizing. Strategically overlooking past developments enables BI proponents to make a grand claim to novelty and innovativeness, so as to mobilize interest and support. While sceptics thus may have rightly observed that the field promotes false anachronisms and clings to inflated rhetoric ('nudging' and 'choice architecture') and 'golden' concepts

('radical incrementalism' and 'Libertarian Paternalism') (Pressman & Wildavsky 1974), the current study may help us to understand *why* this is happening. Given that BI units are still in the business of knowledge brokering and institution building, they need persuasive narratives to dramatize their mission and potential. They (implicitly) understand that they need to 'rebrand' their at most semi-distinctive ideas and methods as uniquely novel in order to gain a license to operate in what is already a crowded institutional space.

In short, this study has shed new light on the question whether there actually is a behavioural state. As juxtaposed in Table 10.1, my observations suggest that rather than an unsubstantial, faddish phenomenon (sceptics) or instead an already powerful, matured state (advocates and critics), BI is more accurately depicted as a fledgling community that is still in the process of institutionalizing. Its emergence has not led to drastic institutional rearrangements within the architecture and processes of government. Its growing presence, however, suggests that it can neither be dismissed as a vacuous fad. Even while still 'under construction', BI is already making inroads in small but non-trivial ways. Groundwork is being laid; the question is whether, when and how 'lift-off' can occur. The best bet for the future of the field's development is on further expansion and maturation rather than stagnation or decline.

**Table 10.1: The state of the behavioural state today**

| | *Trench positions* | *Empirical observations* |
|---|---|---|
| *Is there a behavioural state?* | *Advocates / critics:* The behavioural state is 'here to stay' and already impactful.<br><br>*Sceptics:* The behavioural state is at most a marginal and fuzzy fad, if not mere 'old wine in new casks'. | The behavioural state exists, but it is at the same time still fledgling and in the process of institutionalizing and authorizing itself. Much of everyday practice is about brokering, not directly applying, knowledge. |

## 10.3 How is the behavioural state developing?

Having specified the whereabouts and size of the Dutch BI field, we now proceed by considering what can be said about its developmental trajectory thus far. Which professional boundaries are being formed? How is the field interacting with and fitting into existing governmental institutions? Is it becoming a wholly matured professionalism? Here too, the academic debate has shown little consensus. Advocates and critics have presented BI as a firm, coherent and finalized approach. Sceptics however have taken BI as a fuzzy hype with unset boundaries. They assert

that it consists of 'magic concepts' (Pollitt & Hupe 2011) that may arouse initial enthusiasm but in the end lack practical substance.

My research suggests not one but several developmental directions are being explored. At the frontstage of Dutch practices, the direction seems one of convergence and deepening. BI is presented as a clear-cut set of ideas and methods, corresponding with the narrative of leading Anglo-Saxon role models in the field. This set can be viewed as a 'standardized package' (Fujimura 1992) that combines different elements into one loosely coherent program. Table 10.2 (building on an earlier framework presented in Chapter 6) gives an overview of this program, consisting of five pillars: NBE; individualism; Nudge; Randomized Controlled Trials; and Evidence-Based Policymaking. These five pillars are out there for all to see. They help the BI field to present a persuasive narrative and distinctive program to the outside world. They set the field's boundaries and signal what (ideally) constitutes behavioural policymaking. The implicit assumption behind this frontstage boundary-setting seems to be that practices should, and in principle will cling to these pillars, developing in a convergent fashion.

However, Chapters 5 and 6 have shown how the logic of convergence doesn't depict the field's developmental path accurately, at least with regards to the situation in the Netherlands. In practice, it turns out that the field's professional boundaries are porous. Backstage, behaviour experts are adapting the BI approach to their own local context. This adaptation brings about peripheral but also deeper forms of fragmentation. Deep fragmentation specifically occurs in terms of how members of the behavioural state actually use theories and methods. Fault lines in the field appear between 'purists', who stick to NBE and RCTs, and 'pluralists', who welcome all sorts of other types of knowledge and softer methods.

If we consider the public administration research on policy translation, such backstage fragmentation doesn't come as a surprise. Given that BI is still pioneering and not yet restricted by clear-cut professional boundaries, fragmented development is not only possible but even likely. As the BI agenda travels across the global policy sphere it blends with diverse institutional systems and traditions, resulting in the production of a heterogeneous organizational field with varied adaptations of BI, all tailored to the local political-administrative culture. Moreover, it has to be recognized that Dutch public agencies tend to be relatively autonomous (Schillemans 2012). This absence of significant isomorphic pressures (DiMaggio & Powell 1983) also makes it less likely that the Dutch BI repertoire will eventually converge into a unified collection.

**Table 10.2: The five pillars of Behavioural Insights**

| *Pillar* | *Analytical dimension* | *Description* | *Offered legitimation* |
|---|---|---|---|
| *New Behavioural Economics (NBE)* | Theoretical orientation | Body of knowledge that studies cognitive choice processes, focusing particularly on deviations from rationality | Mainstream economics, assuming *homo economicus,* is insufficiently able to account for real human behaviour. NBE *is.* |
| *Individualism* | Unit of analysis | Perspective that defines and develops policy in relation to separate agents | Traditional policy perspectives do not zoom in deeply enough on individual behaviours, values and choice processes. An individualist perspective *does.* |
| *Nudge* | Instrumental practice | Intervention strategy that redesigns choice environments in subtle ways that tap into aspects of bounded rationality | Traditional policy instrumentation is too often unsuccessful. Nudges, because of their subtle nature and more realistic underpinnings, can be more effective, efficient, and yet less obtrusive. |
| *Randomized Controlled Trials (RCTs)* | Evaluative practice | Evaluation method that tests the causal effects of singular interventions | As individual behaviours are highly complex and context-sensitive, the effectiveness of interventions should be measured. RCTs are the most rigorous way to do so. |
| *Evidence-Based Policymaking (EBP)* | Policy ideology | Policy discourse that promotes the use of scientific evidence and rigorous analysis | Policy decisions are too often made in the context of absent or misused knowledge. EBP helps to make science more instrumental to policy. |

My fieldwork thus reveals a degree of ambivalence about the behavioural state's developmental trajectory. The openly stated ambition is one of converging and deepening, with behaviour experts sticking to their carefully crafted public identity. At the backstage however, the actual direction seems to consist in a more open-ended and divergent uptake of BI elements. BI's shared core, consisting of its five pillars, seems less fixed and directive than it appears at first sight. As true pioneers, behaviour experts freely experiment with a range of 'behavioural' theories and methods. Table 10.3 contrasts these findings with current viewpoints in the behavioural state debate.

**Table 10.3: Development trajectory of the behavioural state**

| | *Trench positions* | *Empirical observations* |
|---|---|---|
| *How is the behavioural state developing?* | *Advocates / critics:* The behavioural state is a fixed and homogeneous project, expanding further in convergent directions.<br><br>*Sceptics:* The behavioural state is incoherent, ill-defined and suffers from critical 'boundary issues' | The behavioural state is still 'under construction'. The field is united by its clear-cut frontstage narrative, emphasizing the value of NBE, RCTs and nudges. Beyond that, the field is loosely fragmented, with varied adaptations of BI all tailored to local institutional context. |

## 10.4 How does the behavioural state see?

While rising to possible future prominence, and developing in ambiguous directions, what insights have we gained in terms of how the behavioural state sees (cf. Scott 1998)? Which types of knowledge does it embrace or instead reject? Which types of actors does it involve in the gathering of knowledge, for instance how does it relate to and interact with the recipients of its policies? What kind of institutional logics and rationalities does it adhere to? I will reflect on the behavioural state's gaze from three conceptual lenses that have also been adopted in the empirical chapters: institutional logics (Chapter 7); deliberative qualities (Chapter 8); and knowledge use and interdisciplinarity (Chapters 8 and 9).

### *Institutional logics*

In the debate, BI has been associated with different types of institutional logics. BI advocates have depicted their approach as 'radical incrementalist' (Halpern 2015), emphasizing their step-by-step working method, designing and testing one nudge at a time. Critics on the other hand have pinpointed the (implicit) rationalist logic in BI, unrealistically framing policy development as a purely rational-analytical affair.

My observations however suggest that the gaze of the behavioural state is more complex than these portrayals. In the Dutch context at least, the behavioural state in fact has a *dual* gaze. Behaviour experts can be *both* rationalist and incrementalist, although at different times and across different stages. In frontstage settings, where they manage their public image (Goffman 1959), they emphasize rationalist values and ambitions. In backstage settings, where the rubber hits the road in direct engagements with policy bureaucrats and stakeholders, behaviour experts seek to reconcile their rationalism with the prevailing institutional logic of incrementalism. A hybrid logic of 'rationalized incrementalism' emerges in which behaviour experts

push their rationalist agenda only to the extent possible given the composition and 'mood' of the environments they find themselves in. By adhering to this more hybrid logic, behaviour experts acknowledge the limits to the full rationalization of policy design, while at the same time trying to remain faithful to their ideal of rigorous evidence-based policy. While doing so, they seem to implicitly understand that their legitimacy is grounded in the 'productive myth' (Boswell 2018) of the rationalist policymaking ideal. Their strategy of seeing with a dual gaze helps them to harness the authorizing power of that myth while also coming to grips with the practical limits thereto.

***Deliberative qualities***

How does the BI movement perceive and interact with its targeted recipients, often citizens? In the debate, critics have claimed that BI has a strong technocratic gaze, using science to 'manage' citizens in a top-down fashion. Advocates on the other hand argue that behavioural policies align with deliberative democracy as they can be, or already are, co-designed, made transparent and publicly debated (John 2018).

My observations suggest that the role of deliberation is more ambivalent, at least in the Dutch BI landscape. On the one hand, there are clear signs of public deliberation about BI. To begin with, behavioural public policy has become part of official policy discourses. Influential government advisory bodies and other think tanks have published reports (e.g. Jonkers & Tiemeijer 2014) discussing the promises and potential pitfalls of BI. News media (e.g. Kleinpaste 20 July 2013) and academic scholars (e.g. Schillemans & De Vries 2016b) have also shown interest in the Dutch BI developments. Behaviour experts themselves too have facilitated public deliberation by organizing communal debates about the merits and (ethical) risks of BI. Furthermore, Chapter 8 has shown that, beyond a broad political and public debate, behaviour experts sporadically also include citizens in particular policy processes. Dutch experts seem slightly more inclined than their Anglo-Saxon counterparts to involve citizens, for instance by seeking their advice during early stages of policy analysis. Dutch experts may also employ behavioural science techniques to achieve deliberative ends, seeking to nudge citizens into participating in political and policy processes.

That said, there is still a strong technocratic undercurrent to the Dutch BI movement. Although citizen benefits and societal well-being lie at the core of its policy narrative, its operational capacity is strongly geared towards applying *scientific* expertise, with only limited scope for citizen input. This observation was confirmed during a focus group, in which participating experts noted that involving citizens during the policy analysis stage was often unfeasible due to time constraints. Citizens are envisaged as

(often unconscious) co-*producers*, not as (explicit and empowered) co-*designers* of BI's policies. Moreover, focus group participants felt that in an overall sense Dutch BI developments took place in a depoliticized fashion. Although they stressed the significance of political dimensions behind the behavioural turn, such dimensions didn't seem to come to the fore through a lively, overt, public political debate.

***Knowledge use and interdisciplinarity***

According to BI's advocates, the heart of BI's knowledge base clearly consists of new behavioural economics and experimental methodology. At the same time, they acknowledge the potential relevance of drawing upon other methods and strands of knowledge: they embrace methodological pluralism and 'genuine interdisciplinarity' (John 2018). Critics, on the other side, have claimed that the field's epistemology is inherently skewed towards experimental methods and to (positivist) knowledge at the interface of psychology and economy. They lament the rise of a 'psychocracy' (Jones et al. 2013) that ignores a whole set of other potentially valuable knowledge perspectives in public problem solving.

My research suggests that in practice the behavioural state's gaze is more hybrid and diverse than its critics assert. Within the BI landscape, there are some *purist* experts whose epistemic practice closely coincides with the theories and methods espoused at the frontstage. Especially some of the key trendsetters, like BIT UK, stick to the narrative in practice, looking at the nature of their projects: mostly RCTs to evaluate the effectiveness of NBE-inspired nudges (e.g. Service et al. 2014). This purist practice corresponds with the prototypical idea of 'behaviourally-tested policymaking' (Lourenço et al. 2016). I have however also observed more *pluralist* experts – at least in the Dutch BI landscape – who embraced a wide range of methods and theories beyond hardcore NBE and sometimes even beyond the domain of behavioural science in general. Such practices indicate that the field's overall still 'bounded interdisciplinarity' (see Chapter 9) is not a given.

This emerging theoretical pluralism is reflected in behaviour experts drawing on a wide range of scientific insights well beyond the narrow domain of NBE. Insights from other branches of psychology and neuroscience in particular have been added to the epistemic package of behaviour experts, often in an ad hoc, context-dependent fashion. To illustrate, an interviewee at the Royal Netherlands Army explained his pragmatic approach to me, adopting new theories and models whenever they intuitively seemed relevant and useful. He mentioned a rich list of scholars of whose work he drew from featuring NBE stalwarts like Kahneman, Thaler and Sunstein but also non-affiliated scholars such as social psychologist Robert Cialdini and neuroscientist Victor Lamme.

A more radical form of pluralism appears to be developing in the field with regards to method use. BI's core espoused method of choice – the RCT – is increasingly being circumvented, downplayed and disputed. A new generation of BI experts deemphasizes RCTs mostly on grounds of feasibility, and resorts to 'satisficing' by turning to easier to apply, softer but still valuable methods and analytical procedures. Behaviour experts in and around government have been adjusting their methodological palette to their local political-administrative circumstances and the constraints these include, even to the point of working with methods – such as professional judgment and intuition – that seem antithetical to the scientific rigour promoted at the frontstage.

In short, this study has shed new light on the behavioural state with regards to how it really sees. Observations have shown how in practice BI's gaze is far more hybrid than its frontstage narrative makes it seem. This hybridity reflects in behaviour experts mixing different institutional logics, combining technocratic and deliberative elements, and partly pluralizing their palette of used theories and methods. Table 10.4 juxtaposes these findings with existing views in the debate.

**Table 10.4: Prisms of the behavioural state**

| | *Trench positions* | *Empirical observations* |
|---|---|---|
| *How does the behavioural state see?* | *Advocates:* The behavioural state follows a 'radical incrementalist' approach. It is deliberative, methodologically pluralist, and 'genuinely interdisciplinary'. | The behavioural state's gaze is hybrid and diverse. This gaze also differs per setting. Frontstage, hardcore rationalism prevails. Backstage, a negotiated 'rationalized incrementalist' logic operates. Furthermore, some BI practices are more technocratic, others more deliberative. Some are more methodologically purist, others more pluralist and critically interdisciplinary. |
| | *Critics:* The behavioural state's gaze is hardcore rationalist, technocratic, and methodologically bounded. | |

Now that we have discussed the first trio of guiding questions for this chapter, let's take stock. So far, I have assessed whether the behavioural state actually exists, how it is developing, and how it sees. I argued that there is indeed a behavioural state, but that it is still in a process of institutionalizing. Furthermore, while at the frontstage this state seems to be developing convergently, in the backstage it actually seems to be hybridizing and diversifying its epistemic prism and practice.

## 10.5 What may become of the behavioural state?

If this is where the Dutch behavioural state is today, it is now time to turn our gaze towards its future. How might its narrative, its activities and its institutionalization evolve? A crucial question here is whether the observed backstage experimentation and deviation will endure (assumption of a diverging future development), or whether practices will eventually move closer towards the frontstage narrative as they mature (assumption of a converging future development). In a way, the Dutch BI landscape's future development comes down to a 'contest' between its frontstage and backstage. This contest may play out in different directions. Let's consider three logically possible, more or less plausible, scenarios in more detail: (I) the frontstage 'wins'; (II) the backstage 'wins'; and (III) it is 'a tie' and the current ambivalent situation endures.

***The 'walking the talk' scenario: hard institutionalization***

A first possible scenario – 'walking the talk' – is that of a relatively hard type of institutionalization, in which backstage practices are pulled up to the frontstage norms. In this scenario, BI's ideal of rigorous and pure NBE-based policy becomes a dominant reality. Such a scenario in which the backstage 'walk' is made compliant with the frontstage 'talk' might seem plausible for various reasons. To begin with, it must be recognized that the field still appears in a positive momentum of growth. As the field will grow in size and resources, it can more easily satisfy the high methodological demands set at the frontstage. A more fully-fledged behavioural state will be much better equipped to anchor its ideas and methods in established policy procedures. This is to some extent already illustrated by resource-high role models like BIT UK, possessing a degree of capacity and autonomy that allows them to run projects in a relatively less constrained fashion. Also, these role models have an important symbolic and example-setting function, so for them there is an extra pressure to 'walk the talk'.

Yet, even if considerable resources and support can be obtained, a pure version of this scenario is not very probable. If nothing else, decades of policy science research have taught us one thing very clearly: there are inherent and considerable limitations to wholly (behaviourally) scientizing the policy process (Lindblom 1959; 1990). As Chapter 7 has demonstrated, behaviour experts can at best succeed in *partially* rationalizing the policy process, resulting in a hybrid 'rationalized incrementalism'. But believing that the behavioural state will one day *fully* and *universally* 'walk its talk' is nothing short of utopian.

***The 'wild west' scenario: soft, disjointed institutionalization***

A second scenario – dubbed 'wild west' – refers to a soft and uncoordinated type of institutionalization trajectory. This scenario holds that BI's backstage practices would continue to broaden and pluralize to such an extent that it becomes fully unclear where BI begins and ends. In this scenario, the diversity of backstage practices becomes so far removed from the frontstage talk that this talk loses its symbolic and example-setting function. Frontstage norms are no longer seen as directive and legitimate forces. There is no perceived need any more for a unified and idealistic frontstage narrative, and instead BI practices can shape their practice as desired, wholly dependent on their own norms, standards and beliefs.

In the near future, it seems relatively unlikely that such a behavioural 'wild west' would occur. In the current political-administrative climate, in which EBP discourse still dominates and in which the related BI trend still receives growing institutional support, there is little reason to believe that behaviour experts will publicly distance themselves from the 'productive myth' (Boswell 2018) of rigorous NBE-based government. This myth seems to provide too many gains for them to detach themselves from it, even if that would result in a wide gap between the front- and backstage.

From a broader historical perspective however it does seem rather more likely that BI's frontstage story would eventually lose its power and legitimacy. Policy paradigms come and go. While the BI trend is currently in vogue, new trends and movements will show up, and new academic insights will claim space and perhaps eclipse some of BI's high ground. Although BI stands in service of a robust ideology of scientization of public policy, it is unlikely that BI will forever remain its most prominent standard bearer moving forward.

***The 'sustained hypocrisy' scenario: soft, loosely connected institutionalization***

In a third scenario – dubbed 'sustained hypocrisy' – BI is institutionalized in a soft, loosely coordinated way, although this softness is not publicly acknowledged by the field itself. The policy talk of the field remains purist and idealistic, assuming and stressing the possibility of a hard type of institutionalization. The policy walk however can deviate. Backstage practices are allowed to develop in diverging directions (provided they still cling to the universal, rationalist frontstage narrative). For instance, experts may conduct macro-level analyses that look beyond cognitive psychological processes, or they may draw on methods much softer than RCTs. By loosening its backstage boundaries, the BI field seems better able to institutionalize across various organizational contexts. Its 'chameleontic' nature makes it more effective at fitting in and with that securing its own future.

'Sustained hypocrisy' is a fairly likely scenario in the near future given that this scenario already seems to best reflect the current Dutch status quo. Knowing that institutions change slowly, if at all (Ansell et al. 2015), the likelihood of this scenario enduring for at least the coming time further increases. Additionally, it ought to be considered that with this scenario the field has thus far had considerable institutional success. Its 'organized hypocrisy' (Brunsson 2007) has been a productive solution to the puzzle of meeting conflicting needs. Hypocrisy first allows for the expression of a grand, idealistic frontstage narrative that provides legitimacy and evokes new hope in better policy, particularly crucial for the field as it is still a new kid on the block. During interviews, field visits and a focus group, this new kid status repeatedly came up as a factor that licensed a strong idealistic narrative. At the same time, hypocrisy also allows for the tailoring of backstage practices so that they fit within established procedures and structures, even when these are not conducive to adoption of the idealistic narrative.

Moreover, 'sustained hypocrisy' seems likely when considering that the BI trend doesn't operate in an institutional vacuum. Instead, BI has to fit in within an already existing policy establishment that is composite of different institutionalized logics (e.g. Bjerregaard 2011) and a pluralism of different traditions and knowledge perspectives. Perspectives such as law, economics and political science have already become deeply anchored in the policy process (Scherpenisse et al. 2016). Even though BI may help bring the NBE perspective more to the forefront, it is implausible that it will overshadow or can distance itself from these other perspectives entirely. The pluralism that is deeply embedded in the policy system creates an overall equilibrated status quo in which different knowledge perspectives may struggle over authority but in the end also merge and/or balance each other out. The 'sustained hypocrisy' scenario allows room for the balancing and merging of different knowledge perspectives in the backstage.

At this stage, it seems that the field's hypocrisy is relatively productive. Loose professional boundaries are being set. Frontrunning Anglo-Saxon units and heroes provide inspiration and a sense of a shared identity, while BI 'backwaters' claim leeway to expand and transcend those boundaries which makes them more adaptable. In the longer run, this hypocrisy may prove difficult to sustain. If the gap between the frontstage talk and backstage walk widens too much, and becomes exposed as such, the field's legitimacy might come to be at risk. The field would then become, in pathological terms, 'schizophrenic', with such a steep frontstage-backstage gap that its overall 'health' would be compromised.

Various responses might be conceivable in reaction to the field's hypocrisy being exposed and challenged. One response might be for the field to 'toughen up' and try harder to be more strictly compliant with the frontstage ideals. It would then seek to move towards the 'walking the talk' scenario. A second, reversed response would be to 'loosen up' and rework its frontstage story into a more 'confessional' narrative that acknowledges the field's porous backstage boundaries. The field then would not need to walk its talk anymore; its walk would *become* the talk. Examples of the frontstage moving closer towards the backstage are as of yet rare but do exist. Two leading role models – BIT UK and the Joint Research Centre of the European Commission – have recently published reports with fairly atypical views, explicitly pushing for the use of non-experimental methods in the field (Van Bavel & Dessart 2018) or reflecting on the field's tacit rationalist assumptions (Hallsworth et al. 2018). A third, more extreme response would be for the field to delegitimize its purist story altogether, instead celebrating the deviant, experimental and pluralist nature of its practices. At the frontstage level, the field would then disjoin, bringing it back to the 'wild west' scenario. These abovementioned responses are however fragile and may fall short. The first response, trying to walk the talk, can be unrealistic in an incrementalist policy environment full of constraints. The second and third response also are unlikely to provide the field with enduring legitimacy, as both suffer from the lack of a purist narrative that provides direction, unity and appearance of competence – which inevitably will impact the field's institutionalization trajectory. In a later future, when neither hypocrisy nor any other scenario would work out, a fourth 'response' might then be that both the BI's frontstage *and* backstage would fall into demise.

Table 10.5 shortly recapitulates the three future development scenarios, while considering their associated frontstage-backstage dynamic and their likeliness of happening.

**Table 10.5: Scenarios for the future of the behavioural state**

| *Development scenario* | *Frontstage-backstage dynamic* | *Description* | *Likeliness* |
|---|---|---|---|
| *'Walking the talk'* | Frontstage makes backstage compliant | Hard institutionalization. The ideal of rigorous and pure NBE-based policy is realized in practice. | Likely in the case of resource-high role models, but less likely in an overall sense due to inevitable limitations to wholly (behaviourally) scientizing the policy process. |
| *'Wild west'* | Backstage overrules and disjoins frontstage | Extremely soft, disconnected institutionalization. Backstage practices develop in wholly diverging directions, while disempowering the frontstage narrative. | Less likely at this time given the legitimacy provided by the frontstage narrative. But possible in a later future considering that attempts at (behavioural) scientization have historically come and gone. |
| *'Sustained hypocrisy'* | Organized split between frontstage and backstage | Soft, loosely coordinated institutionalization. Backstage practices develop in slightly diverging directions while still supporting and reproducing the frontstage narrative. | Fairly likely for the moment. Still in the process of institutionalizing, the field is experiencing a need for a grand, promising, idealistic narrative, while equally there is a need to adapt, compromise and satisfice practices given practical limitations and local demands. As long as it remains unexposed, organized hypocrisy helps to meet these conflicting needs. |

## 10.6 Should there be a behavioural state?

Moving from forward mapping to normative reflection, the next question to be discussed is whether a behavioural state is 'good' and desirable for our society. I would contend that the answer depends on how the behavioural state will evolve. Below, I will evaluate and compare the three development scenarios sketched above, as they all come with their own institutional arrangements, practices and social impacts. What are their respective strengths/opportunities, but also what unique risks/weaknesses do they come with?

***Appraising the 'walking the talk' scenario***

In this scenario, a fully-fledged behavioural state has developed in a convergent way, conforming to the pillars of the field. The main benefit hereof would be that NBE insights and methods are most purely and comprehensively imported in established policy procedures. BI would then actually fulfil its mission, broadly interpreted as the psychological enrichment of policy practice, particularly with NBE knowledge. The resulting aggregation of produced analyses, interventions and evaluations (throughputs) might eventually reach the point of a 'critical mass' of behavioural public policy which would actually deliver on the promised large-scale social change (outcomes) across policy domains.

Yet, when BI would become more able to actually walk its talk, which is not wholly unlikely in those contexts where BI is in popular demand, a potential 'dark side' might grow with it. The main risk is that the field would go too far in seeking to capitalize on its strengths and reap its full benefits, obsessively paying attention to cognitive psychological dimensions at the micro-level of public policy issues. This risk of such 'epistemic tunnel vision' in which policy discourse becomes centred on people's cognitive limitations, biases and errors at the micro-level might also be coined the 'microfication' (Jankowski 2007) or 'overpsychologization' (Jankowski 2007; Timmer 1975) of policy. Problematic about such overpsychologization is that it can overlook the influence of non-psychological dimensions and the possibilities of intervening at the meso- and macro-level. Also, it may implicitly support a discourse that holds 'flawed' individual decision-makers overly accountable for social issues.

Indications that overpsychologization is a real risk for the current BI movement can already be observed. In the area of poverty and debt policy, for example, behaviour experts have been increasingly concentrating on analysing and targeting cognitive decision-making 'flaws' of debt holders and poor people. For instance, the recent BIT UK report (Gandy et al. 2016) recounts new insights regarding psychological factors that shape financial behaviour. While in and of itself insightful, such an analysis implicitly shifts the attention away from other relevant factors, such as existing economic inequalities, neighbourhood deterioration and 'the culture of poverty' (Lewis 1968). Tunnel vision also loomed large in the Flint water contamination case discussed in Chapter 9, where the US Social and Behavioral Sciences Team initially started exploring the psychological mechanisms that could be leveraged to steer individual citizens towards simple behaviours (e.g. washing their hands), only to discover that this approach failed to address the more urgent, deeper-lying problems of lacking institutional trust and economic inequalities. The risk of the psychological perspective becoming overbearing also looms in the Dutch BI landscape, whether we for instance consider the 'School Canteen Brigade Officers' leveraging psychological

mechanisms to improve healthy food choices of schoolchildren, or the local behaviour experts experimenting with psychological techniques to induce citizens to repay their debts faster. Should these types of BI projects multiply they may add up to a combined trend of radical policy individualization and psychologization, in which an increasingly dominant psychological gaze gives rise to an individualist policy drive to target, analyse and responsibilize citizens in an attempt to tackle social issues (e.g. around health, personal finance and climate change) while responsibilities of for instance corporate businesses become increasingly eclipsed.

Another risk of a fully-fledged behavioural state would be its eagerness to intervene in public space. Although BI generally claims to respect and even promote freedom, it tends to interpret that concept rather narrowly (Vugts et al. 2018). Mostly, 'freedom' is framed as a technical matter of whether or not an intervention leaves room for an 'opt-out', or as an issue of whether an intervention promotes policy recipients' own interests as judged by themselves. In the BI narrative, the need to protect positive liberties of citizens ('freedom *to*') tends to be emphasized, while negative liberties ('freedom *from*') tend to be devalued, problematized and/or narrowly defined (Berlin 1958, also see McLaughlin 2016). The dominant reasoning is that *some* kind of choice architecture is inevitable and that behavioural influence simply is ubiquitous, causing the field to automatically problematize the alleged negative liberties of citizens and instead shifting the focus to supporting their positive liberties through more consciously designed public choice architectures.

This reasoning however lacks nuance. Even if behavioural influence as a general phenomenon is inevitable, there are still important moral choices to be made in terms of how deep state interference in citizens' everyday lives should be. Governments must decide how much of an effort they put into analysing and subtly redesigning individual citizens' behaviours. They must form a view of how much they value citizens' negative liberties - not just narrowly interpreting them in terms of whether there are technical 'opt out' possibilities. In other words, they must decide (and defend) the appropriate depth of the state's influence in designing the *context* and *process* in which citizens make their choices (Feitsma & Schillemans 2015). A fully-fledged behavioural state however is unlikely to consider such choices as long as the current BI narrative, which largely disregards such matters, continues to guide its practices. Rather, BI's interventionist gaze is likely to intensify, justified by the implicit claim that it is positive liberties that matter most. In other words, a fully-fledged behavioural state might become 'greedy' (Trommel 2009), wholly geared towards the detailed analysis and architecture of public space without considering the moral impact of its intensified presence in citizens' everyday lives.

***Appraising the 'wild west' scenario***

In this scenario, the institutionalization of BI occurs in a wholly disjointed fashion, with a proliferation of practices that have become detached from the field's espoused pillars. This has the advantage of allowing ample space for polycentric experimentation and learning, as localized BI practices try out alternative ideas and methods. Potentially, this development of BI would constitute a productive form of 'experimentalist governance' (Sabel & Zeitlin 2012). Moreover, this scenario would also create space for critical interdisciplinary engagement, avoiding it being captured by adherents of its early pillars like NBE, RCTs and nudges. Exploring new territory through an array of perspectives might help make the field's epistemic practices more sophisticated and better tailored to the realities of complex policy issues.

At the same time, when a directive purist frontstage is lacking, the BI field will be less able to import its core knowledge foundation (i.e. NBE) in a pure and comprehensive fashion. BI's original promise will then not likely be fulfilled. Another downside of the 'wild west' scenario is that the field's identity may become blurred by elemental 'boundary issues'. In Chapter 6, for example, the steep fault lines between methodological purists and pluralists came to the fore. Can the field convincingly include both purists wedded entirely to RCTs and pluralists who base their policy designs on many methods, including rough estimation and practical intuition? Such fault lines raise puzzling questions about where BI begins and ends. As BI practices both proliferate and diversify in ways not held together by an integrative frontstage story, it will likely become more difficult to recognize BI's distinctive competence and value proposition. An analogy might be made to the biblical story (Genesis 11:9) about the Tower of Babel. This story depicts Babylonians trying to build a tower to reach God, eventually being hindered in their attempts as God creates a confusion of languages between them. Similarly, the 'wild west' scenario's risk is that when widely different 'BI languages' are formed and allowed to co-exist, it becomes less and less possible for experts to collaborate and collectively legitimize themselves. The risk, or 'tragic side', of the field falling into a 'Babylonic collapse' then looms large.

***Appraising the 'sustained hypocrisy' scenario***

In this scenario, which best reflects the current Dutch status quo, BI is institutionalized in a soft fashion, with a coherent and somewhat directive frontstage but also room for backstage experimentation. BI experts loosely adhere to the field's pillars-dominated frontstage story and yet also when needed deviate from it in their tailor-made backstage practices. The likely result of this scenario is that at least partial progress is made in importing BI's foundational ideas and methods into policy, while the behavioural state is at the same time prevented from becoming overbearing. This scenario thus has the best prospect for ensuring BI leads to sustainable epistemic

enrichment and better public policy (its chief promise), without it in practice suffering from epistemic 'overkill' and turning into a bulwark of technocracy and psychocracy (its potential 'dark side').

A drawback of this scenario is that the field will probably prove unable to fully deliver on its original founding myth. The more behaviour experts adapt and pluralize their practices in the backstage, the less likely they are to be able to import their distinctive ideas and methods. A more urgent risk of the 'sustained hypocrisy' scenario is, however, that it still can come with a relentlessly psychocratic policy *talk* (vs. walk). The 'dark side' of an overpsychologized *narrative* is actually especially likely in this scenario, considering that organized hypocrisy works through a logic of compensation between the domains of talking, deciding and acting (Brunsson 2007). *Because* the field can adapt and pluralize at the backstage, it can remain hardcore purist at the frontstage. *Because* its walk is soft, its talk can be rigid. One might view an uncompromised overpsychologized policy talk as relatively unproblematic: it is not the talk but the walk that matters. However, this ignores the performative effects of policy narratives: the talk can come to shape the practice, informing it in tacit but nonetheless powerful ways. Heavily grounded in cognitive psychology, BI's policy talk may contribute to an overly individualized political discourse that excessively tends to frame citizens as primary problem sources, policy targets and responsibility bearers (Shove 2010). Policy problems are then automatically framed as 'self-responsible individuals behaving in undesired ways', and policy solutions become reduced to 'individual behaviour change interventions' – importantly, not for analytical reasons per se but rather because such framings are compliant with the dominant political zeitgeist about how social responsibilities ought to be allocated.

***Weighing the three scenarios***

Table 10.6 summarizes the strengths/opportunities and risks/weaknesses of the three development scenarios. A pivotal question is whether the behavioural state can be developed in a 'tamed' fashion, optimally capitalizing on its core strengths (i.e. psychological enrichment) while dodging a potential 'dark side' (i.e. overpsychologization) or 'tragic side' (i.e. loss of legitimacy). The 'sustained hypocrisy' scenario, which best reflects the current Dutch status quo, seems to have the best chances at this taming prospect: this scenario features a mildly directive frontstage yet also some room for backstage deviation and experimentation. As such, a seemingly desirable balance is found between BI's need to push its NBE perspective as purely and strongly as possible, and society's need to have that perspective be an epistemic enrichment, not a replacement. Nonetheless, a perfect taming seems unlikely as this scenario still may give rise to an overpsychologized policy talk in which citizens are indubitably framed as primary responsibility bearers and

individual behaviour change strategies as default solutions.

**Table 10.6: Appraising the futures of the behavioural state**

| *Scenario* | *Strengths / Opportunities* | *Risks / Weaknesses* |
|---|---|---|
| *'Walking the talk'* | BI's insights and methods are most purely and comprehensively imported in policy. BI's essence and identity remain fully intact. | 'Dark side' of an overpsychologized talk *and* walk, and threat to negative liberties. |
| *'Wild west'* | Plenty of possibilities for deep policy learning, experimentation, and critical interdisciplinary engagement. | 'Tragic side' of BI losing its identity, falling into a 'Babylonic collapse'. |
| *'Sustained hypocrisy'* | Best chances of developing the BI field in a 'tamed' fashion. BI's foundational insights and methods can be partially imported without the field becoming overbearing. | 'Dark side' of an overpsychologized policy talk. |

## 10.7 How should the behavioural state be studied?

What does this study suggest by way of urgent and productive avenues of future inquiry into the BI movement and its attempt to nestle itself within the state? I see a vital role for two types of research agendas in particular: practice-oriented/ ethnographic and critical-reflexive/interdisciplinary approaches. To a modest extent these agendas already exist, and this study might be seen as an exploratory example hereof. Yet, I would contend that they can be deepened and expanded.

***Practice-oriented and ethnographic approaches***

Chapter 1 has argued that the current behavioural state debate tends to lack a proper empirical foundation. The BI practices and their actual policy implications tend to be misrepresented by either one of the camps within the debate. The advocates' grand promises seem inflated, the critics' dystopian accounts overblown, and the sceptics' brush-offs too simplistic. *Practice-oriented* approaches can help to move beyond the current trench warfare. Following Lepenies and Malecka (2018), I call for a 'practices-turn' in research on BI. Practice-oriented approaches move away from questions about what behavioural public policymaking *should* look like or *could* be and instead capture how it currently is practiced. They dive deeper into what is actually happening in the field, and what the people in it are actually doing. The importance of studying BI practices continues to be important considering that the field is still 'under construction', and that, as has been argued above, different future development trajectories are possible and plausible. Scholars thus ought to keep

monitoring major trends in the field, studying how BI is practiced across institutional and geographical boundaries. Furthermore, they ought to capture these trends in a way that does justice to the richness of the global BI landscape. It remains important to study beyond the 'usual suspects' (i.e. the BITs in the Anglosphere). This study has made a small step in this by exploring the little explored Dutch BI landscape. However, much other landscapes remain relatively opaque (e.g. Japan, Singapore, Qatar). By further mapping out how BI is being institutionalized in various contexts, we can more clearly see the 'varieties of behavioural expertise' (Strassheim & Korinek 2015) that are emerging. Chapter 6 has made a small step in this direction by exploring key methodological contingencies in the global BI landscape. But much deeper insights could be extracted, especially regarding how BI practices are shaped by their local institutional context.

As part of the practices-turn, *ethnographic* approaches seem of particular importance as they help capture the backstage of behavioural practices (Van Hulst 2008). Backstage access is needed given the observed frontstage-backstage differences in the field. BI's backstage walk thus is likely to provide valuable information that cannot be retrieved from its frontstage talk. Yet, ethnographic research on BI practices remains scarce, and, this study included, also still relatively short term and shallow when compared to some of the more traditional ethnographic studies based on years of deeply immersed participant observation. More fully immersed ethnographic work should be conducted, which will prove especially helpful in observing and analysing (changing) frontstage-backstage dynamics in the field.

***Critical-reflexive and interdisciplinary approaches***

This study has suggested that the behavioural state brings potentially powerful gains in terms of its psychological enrichment of policymaking. At the same time, in an excessive form it also comes with risks and a 'dark side', i.e. an overpsychologized policy talk and/or walk. Therefore, the behavioural state's ongoing developments need to be monitored and assessed in relation to these risks. Critical normative scrutiny of this state's legitimacy remains necessary, considering that there is little reason to expect that BI's risk of overpsychologization will fade soon. Still seeking to institutionalize, the field is likely to continue to draw legitimacy from preaching a strict, purist, heavily psychologized message. Moreover, BI ought to be seen as part of a broader political zeitgeist in which cognitive psychological and neurological sciences are in demand, reflected in the recent emergence of bodies of thought such as *neurocapitalism, neuropower* and *neuroliberalism* (Rentea 2016; Whitehead et al. 2018). So long as overpsychologization remains a plausible risk we need scholars to 'speak truth to power' (Wildavsky 1979), calling the BI field on its assumptions about 'good'

policymaking and the 'logical' social contract, and critically scrutinizing the impacts thereof on society.

Importantly, such critical scrutiny should be grounded in an empirical understanding of actual BI practice. Only then, critical scholars can function as a legitimate watchdog for the potential 'dark side' of a maturing behavioural state. Examples of scholars who both critically *and* thoroughly empirically evaluate the behavioural state's legitimacy exist but are rare (e.g. Strassheim et al. 2015; Lepenies & Malecka 2018; Mulderrig 2018). Also, they seem to only reach a small, and exclusively academic audience. Non-critical, confirmative appraisals on the other hand seem to have a much wider reach and also enjoy much stronger institutional support. For instance, consider Sunstein's *Why Nudge* (2014), which at the moment of writing has been cited 505 times, whereas Mulderrig's (2018) article 'Multimodal strategies of emotional governance: a critical analysis of 'nudge' tactics in health policy' in *Critical Discourse Studies* has been cited only 2 times thus far. To ensure that the more critical voice, and its urgently needed watchdog function, is not overshadowed in the debate, I call for more *critical-reflexive* research approaches. Such approaches critically assess underlying assumptions and values in the BI field, pinpoint potential drawbacks and risks, and explore new desirable paths forward.

This plea can be seen as part of a broader call for more critical-reflexive (and also more public) research in the field of public administration (Karré et al. 2017; Van Putten & Trommel 2018). More and more, public administration (PA) scholars are producing knowledge that either stands in the service of highly specialized theory development and advances in methodological rigour ('professional PA'), or in the service of the instrumental contribution to particular policy issues ('policy PA') (Karré et al. 2017). The focus on these types of knowledge production reflect enduring trends of specialization and instrumentalization within academia. While these trends bring important gains for both the scholarly and policy community, the risk is that they overshadow critical-reflexive approaches ('critical PA') in the field. Unlike 'professional PA' and 'policy PA', 'critical PA' dares to ask 'bigger', more holistic questions, arising from a desire to address urgent societal developments. It dares to be normative and future-oriented, exploring how societal trends will evolve and whether we should deem these desirable (Flyvbjerg 2001).

*Interdisciplinary* research approaches can be of particular help in strengthening a critical-reflexive research agenda. These approaches infuse and confront the BI field with insights and associated debates from other scientific disciplines. Such confrontation helps to keep BI's epistemic gaze broadly and openly oriented, guarding it against the risk of overpsychologization. The body of critical

interdisciplinary research is however as of yet fairly limited (e.g. Whitehead et al. 2017). This study has made a start in exploring new interdisciplinary territories. It has brought in relatively distant debates (e.g. on rationalism vs. incrementalism in policy design) and disciplines (e.g. sociology and human geography). Yet, a vast territory of interdisciplinary engagement remains unexplored. As a first attempt to set the research agenda, Chapter 9 (see Table 9.1) has sketched an outline for critical interdisciplinary analysis. Preferably, such analysis should not be carried out from an 'arm chair' position, from a comfortable distance stating which knowledge perspectives are missing in standard BI practice. Scholars ought to adopt a more constructive role, engaging and collectively 'puzzling' with the field as a productive route towards transforming the field in desirable ways.

# References

@faisal_naru. 8 August 2018. Tweet. Retrieved from: https://twitter.com/faisal_naru/status/1027162896340578304

@Marielle1972. 21 June 2018. Tweet. Retrieved from: https://twitter.com/NOS/status/1009460884865994756

Abbott, A. (1988).*The Systems of Professions: An Essay on the Division of Expert Labor*. Chicago: University of Chicago Press.

Alemanno, A. & Sibony, A. (2015). *Nudge and the Law: A European Perspective*, Oxford and Portland: Hart Publishing.

Akerlof, G. A. & Shiller, R. J. (2009). *Animal spirits: How human psychology drives the economy, and why it matters for global capitalism.* Princeton: Princeton University Press.

Ames, P., & Hiscox, M. (2016). *Guide to Developing Behavioural Interventions for Randomised Controlled Trials*. Canberra, Australia: Commonwealth of Australia

Ansell, C., Boin, A. & Farjoun, M. (2015). Dynamic Conservatism: How Institutions Change to Remain the Same. *Research in the Sociology of Organizations* 44: 89–119

Ariely, D. (2008). *Predictably irrational: The hidden forces that shape our decisions.* New York: Harper.

Arno, A. & Thomas, S. (2016). The efficacy of nudge theory strategies in influencing adult dietary behaviour: a systematic review and meta-analysis. *BMC public health*, 16: 676.

Agar, M. H. (1996). The professional stranger: An informal introduction to ethnography. Bingley: Emerald Publishing Limited

Barber, B. R. (2013). *If mayors ruled the world. Dysfunctional nations, rising cities.* New Haven: 2013.

Baldwin, R. (2014). From regulation to behaviour change: Giving nudge the third degree. *The Modern Law Review*, *77*(6): 831-857.

Ball, S. & Feitsma, J. N. P. (2018). Nudge in Translation. Comparing Behavioural Insights Practices in Australia and the Netherlands. EGPA 2018 Conference Paper.

Ball, S., M. Hiscox, & Oliver, T. (2017). Starting a Behavioural Insights Team: Three Lessons from the Behavioural Economics Team of the Australian Government." *Journal of Behavioral Economics for Policy,* 1: 21-26.

Bauman, Z. (2000). *Liquid Modernity.* Cambridge: Polity Press.

BBC News. (4 December 2017). Wales' organ donor opt-out law has not increased donors. Retrieved from: https://www.bbc.co.uk/news/uk-wales-42213813

Beck, G. (28 September 2010). Glenn Beck: From nudge to shove. Fox News. Retrieved from https://www.foxnews.com/story/glenn-beck-from-nudge-to-shove.

Beck, U. (1992). Risk society: Towards a new modernity. London/Newbury Park: Sage Publications.

Behavioural Public Policy Website (2018). Retrieved from: https://www.cambridge.org/core/journals/behavioural-public-policy

Behavioral Scientist. (4 September 2018). Nudge Turns 10: A Special Issue on Behavioral Science in Public Policy. Behavioral Scientist. Retrieved from: http://behavioralscientist.org/nudge-turns-10-a-special-issue-on-behavioral-science-in-public-policy/

Bemelmans-Videc, M., Rist, R. & Vedung, E. (2003). *Carrots, Sticks, and Sermons: Policy Instruments and Their Evaluation.* New Jersey: Transaction Publishers New Brunswick.

Bendor, J. (2015). Incrementalism: Dead yet Flourishing. *Public Administration Review, 75*(2): 194-205.

Benson, B. (1 September 2016). *Cognitive bias cheat sheet.* Retrieved from: https://betterhumans.coach.me/cognitive-bias-cheat-sheet-55a472476b18

BETA. (2018a). *Behavioural Insights for Policy: Case Studies from Around Australia*. Canberra, Australia: Commonwealth of Australia.

BETA. (2018b). *Saying more with less: Simplifying energy fact sheets*. Canberra, Australia: Commonwealth of Australia.

BETA. (n.d.-a). About. Retrieved from: http://behaviouraleconomics.pmc.gov.au/about

BETA. (n.d.-b). BETA Registered Trials. Retrieved from: https://www.pmc.gov.au/domestic-policy/behavioural-economics/beta-registered-trials

BETA & Behavioural Economics and Research Team. (2018). *Nudge vs Superbugs: A behavioural economics trial to reduce the overprescribing of antibiotics*. Canberra, Australia: Commonwealth of Australia.

Beuving, J., & de Vries, G. (2014). *Doing qualitative research.* Amsterdam: Amsterdam University Press.

Berlin, I. (1958). *Two Concepts of Liberty.* An Inaugural Lecture delivered before the University of Oxford on 31 October 1958. Oxford: Oxford University Press.

Bernays, E. L. (1928). *Propaganda.* New York: Horace Liveright Inc.

Biesta, G. (2007). Why 'what works' won't work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory, 57*(1): 1-22.

BIG'R (Behavioural Insights Group Rotterdam) Website (2018). Retrieved from: https://www.bigrotterdam.nl/

BIN NL. (2017a). *Rijk aan gedragsinzichten: Editie 2017.* Den Haag: Ministerie van Economische Zaken en Klimaat.

BIN NL. (2017b). *Zeven Behavioural Insights Tools.* Retrieved from: https://www.communicatierijk.nl/documenten/publicaties/2017/11/23/tooloverzicht-bin-nl

BIN NL. (2018). *A Wealth of Behavioural Insights: 2017 Edition*. The Hague: Ministry of Economic Affairs and Climate Policy.

BIT UK. (n.d.-a). BI Ventures. Retrieved from http://www.behaviouralinsights.co.uk/ventures/

BIT UK. (n.d.-b). Who we are. Retrieved from http://www.behaviouralinsights.co.uk/about-us/

BIT UK & Department of Energy and Climate Change. (2014). *Evaluation of the Department for Energy and Climate Change and John Lewis Energy Labelling Trial*. London, UK.

Bjerregaard, T. (2011). Co-Existing Institutional Logics and Agency among Top-Level Public Servants: A Praxeological Approach. *Journal of Management and Organization, 17*(2): 194-209.

Bogliacino, F., Codagnone, C., & Veltri, G. A. (2016). An introduction to the special issue on "the behavioural turn in public policy: new evidence from experiments". *Economia Politica, 33*(3): 323-332.

Bonell, C., McKee, M., Fletcher, A., Haines, A. & Wilkinson, P. (2011a). Nudge smudge: UK Government misrepresents "nudge". *The Lancet, 377*(9784): 2158-2159.

Bonell, C., McKee, M., Fletcher, A., Wilkinson, P. & Haines, A. (2011b). One nudge forward, two steps back. *BMJ, 342:* 241-242.

Box, R. C. (2005). *Critical Social Theory in Public Administration.* New York/London: M.E. Sharpe.

Boogers, M. J. G. J. A. & Schaap, L. (2007). Bestuurskracht: wat moeten we ermee? Een stand van zaken. *Bestuurskunde*, 2: 5-14.

Boll, K. & Rhodes, R. A. W. (2015). Excursions in administrative ethnography. *Journal of Organizational Ethnography, 4*(2). doi: 10.1108/JOE-05-2015-0013

Boswell, J. (2017). What Makes Evidence-Based Policy Making Such a Useful Myth? The Case of NICE Guidance on Bariatric Surgery in the United Kingdom. *Governance, 31:* 199-214.

Bovens, L. (2008). The ethics of nudge. In Grüne-Yanoff, T. & S.O. Hansson (Eds.), *Preference Change: Approaches from Philosophy* (207–219). Berlin/New York: Springer.

Bowen, G. (2008). Naturalistic inquiry and the saturation concept: A research note. *Qualitative Research, 8*(1): 137-152.

Brunsson, N. (2007). *The Organization of Hypocrisy: Talk, Decisions and Actions in Organizations*. Copenhagen: Copenhagen Business School Press.

Bucher, T., Collins, C., Rollo, M. E., McCaffrey, T. A., De Vlieger, N., Van der Bend, D., Truby, H. & Perez-Cueto, F. J. A. (2016). Nudging consumers towards healthier choices: a systematic review of positional influences on food choice. *British Journal of Nutrition*, *115*(12): 2252-2263.

Button, M. E. (2018). Bounded Rationality without Bounded Democracy: Nudges, Democratic Citizenship, and Pathways for Building Civic Capacity. *Perspectives on Politics*, *16*(4): 1034-1052.

Cairney, P. (2016). *The Politics of Evidence-Based Policy Making.* London: Palgrave Macmillan.

Camerer, C. F., Loewenstein, G., & Rabin, M. (Eds.). (2011). *Advances In Behavioral Economics.* Princeton: Princeton University Press.

Campbell, I. D. (2017). Cleverer than command? Review of David Halpern, inside the nudge unit. *Social and Legal Studies, 26*(1): 111-126.

Cairney, P. (2017). Evidence-Based Best Practice Is More Political than It Looks: A Case Study of the 'Scottish Approach'. *Evidence & Policy, 13*(3): 499-515.

Carter, E.D. (2015). Making the Blue Zones: Neoliberalism and nudges in public health promotion. *Social Science & Medicine, 133*: 374-382.

Cappellaro, G. (2017). Ethnography in public management research: A systematic review and future directions. *International Public Management Journal*, *20*(1): 14-48.

Chew, S., Armstrong, N. & Martin, G. (2013). Institutionalising knowledge brokering as a sustainable knowledge translation solution in healthcare: How can it work in practice? *Evidence & Policy, 9*(3): 335-351.

Christensen, T. & Lægreid, P. (2002). New public management: Puzzles of democracy and the influence of citizens. *Journal of Political Philosophy*, *10*(3), 267-295.

Cialdini, R. B. (1984). *Influence: The Psychology of Persuasion*. New York: HarperBusiness.

Conly, S. (2012). Against autonomy. Cambridge: Cambridge University Press.

Clarence, E. (2002). Technocracy reinvented: The new evidence based policy movement. *Public Policy and Administration, 17*(3): 1-11.

Cohen, M. D., March, J. G. & Olsen, J. P. (1972). A Garbage Can Model of Organizational Choice. *Administrative Science Quarterly, 17*(1): 1-25.

Collier, D. (1993). The Comparative Method. In A. W. Finifter (Ed.), *Political Science: The State of the Discipline II* (pp. 105-119). Washington: American Political Science Association.

Davies, W. (2014). *The Limits of Neoliberalism: Authority, Sovereignty, and the Logics of Competition.* London: Sage.

Davies, H. T. & Nutley, S. M. (Eds.). (2000). *What works?: Evidence-based policy and practice in public services*. Bristol: Policy Press.

Davies, T. (1997). *Humanism: The new critical idiom.* London/New York: Routledge.

Deaton, A. & Cartwright, N. (2018). Understanding and Misunderstanding Randomized Controlled Trials*. Social Science & Medicine, 210:* 2-21.

De Jong, S. & Rusman, F. (2 March 2015). Denkt u aan uw aangifte? Dank! Liza en Joyce. NRC.

De Jonge, P., Zeelenberg, M. Verlegh, P. W. (2018). Putting the public back in behavioral public policy. *Behavioural Public Policy*, *2*(2): 218-226.

Dijksterhuis & Van Baaren. (2015). Menukaart Participatie op het gebied van zwerfafval stimuleren en vasthouden. Retrieved from: http://www.kenniswijzerzwerfafval.nl/download_document/844.

DiMaggio, P. & Powell, W. (1983). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review, 48*(2): 147-160.

Dolan, P., Hallsworth, M., Halpern, D., King, D. & Vlaev, I. (2010). *MINDSPACE: Influencing Behaviour Through Public Policy*. London: Cabinet Office and Institute for Government.

Dolowitz, D., & Marsh, D. (1996). Who learns what from whom: A review of the policy transfer literature. *Political Studies, 44*(2): 343-357.

Dorren, L. (2015). *Principes in uitvoering.* Master's Thesis. Utrecht School of Governance.

Dror, Y. (1964). 'Muddling Through' - 'Science' or Inertia? *Public Administration Review 24*: 153-157.

Dunt, I. (5 February 2014). Nudge nudge, say no more. Brits' minds will be controlled without us knowing it. The Guardian.

Dunleavy, P., & Hood, C. (1994). From old public administration to new public management. *Public money & management*, *14*(3): 9-16.

Dunleavy, P., Margetts, H., Bastow, S. & Tinkler, J. (2006). New public management is dead – long live digital-era governance. *Journal of public administration research and theory*, *16*(3): 467-494.

Easton, S. (28 November 2016). Experimental trials: a BETA way of making policy. The Mandarin.

Einfeld, C. (2018). Nudge and evidence based policy: fertile ground. *Evidence & Policy,* advance online publication. doi: 10.1332/174426418X15314036559759

Esmark, A. (2019). Communicative governance at work: how choice architects nudge citizens towards health, wealth and happiness in the information age. *Public Management Review*, *21*(1): 138-158.

Esterling, K. M. (2009). *The political economy of expertise: Information and efficiency in American national politics.* Ann Arbor: University of Michigan Press.

Etzioni, A. (1967). Mixed-Scanning: A 'Third' Approach to Decision-Making. *Public Administration Review, 27*: 385-392.

Faulconbridge, J. R. & Muzio, D. (2012). Professions in a globalizing world: towards a transnational sociology of the professions. *International Sociology*, *27*(1): 136-152.

Feitsma, J. N. P. (2016). Meer dan een nudge: Gedragsexperts bij de Nederlandse overheid. *Bestuurskunde, 25*(3): 24-32.

Feitsma, J. N. P. (2018a). The behavioural state: critical observations on technocracy and psychocracy. *Policy Science, 51,* 387-410.

Feitsma, J. N. P. (2018b). 'Rationalized Incrementalism'. How behavior experts in government negotiate institutional logics'. *Critical Policy Studies,* advance online publication. doi: 10.1080/19460171.2018.1557067

Feitsma, J. N. P. (2018c). De makelaarsmachinerie achter gedragsbeleid. *Christen Democratische Verkenningen,* 2: 74-79.

Feitsma, J. N. P. (2019). Brokering behaviour change: The work of Behavioural Insights experts in government. *Policy & Politics, 47*(1): 37-56.

Feitsma, J. N. P. & Schillemans, T. (2015). Gelukspolitiek en de gedragswetenschappelijke uitdaging. *Liberaal Reveil, 56*(2): 102-105.

Feitsma, J. N. P. & Schillemans, T. (2019). Behaviour experts in government: From newcomers to professionals? In H. Strassheim & S. Beck (Eds.), *Handbook of Behavioural Change and Public Policy.* Cheltenham: Edward Elgar Publishing.

Fischer, F., Torgerson, D., Durnova, A. & Orsini, M. (2015). Introduction to critical policy studies. In F. Fischer, D. Torgerson, A. Durnova, & M. Orsini (Eds.), *Handbook of critical policy studies* (pp. 1-24). Cheltenham: Edward Elgar Publishing.

Fitzpatrick, M. (2011). Mad men take over the coalition. *British Journal of General Practice*, *61*(582): 71-71.

Fleming, J., & Rhodes, R. A. W. (2018). Can experience be evidence? Craft knowledge and evidence-based policing. *Policy & Politics, 46*(1): 3-26.

Flyvbjerg, B. (2001). *Making Social Science Matter.* Cambridge: Cambridge University Press.

Forester, J. (1993). Critical Theory, Public Policy, and Planning Practice. Albany: SUNY Press.

Freeman, R. (2007). Epistemological bricolage: How practitioners make sense of learning. *Administration & Society, 39*(4): 476-496.

Freeman, R. (2009). What is 'translation'? *Evidence & Policy, 5*(4): 429-447.

Freeman, R. & Maybin, J. (2011). Documents, Practices and Policy. *Evidence & Policy, 7*(2): 155-170.

Freidson, E. (2001). *Professionalism, the Third Logic. On the Practice of Knowledge*. Cambridge: Polity.

Frissen, P. H. A. (2013). *De Fatale Staat.* Alkmaar: Van Gennep.

Frodema, R. & Mitchum, C. (2007). New Direction in Interdisciplinary: Broad, Deep, Critical. *Bulletin of Science, Technology & Society*: 506-514.

Foucault, M. (2007). *Security, Territory, Population: Lectures at the Collège de France, 1977-78.* Edited by Michel Senellart. Translated by Graham Burchell. London: Palgrave Macmillan.

Fujimura, J. (1992). Crafting science: Standardized packages, boundary objects, and "Translation". In A. Pickering (Ed.), *Science as Practice and Culture*. Chicago: University of Chicago Press.

Furedi, F. (2011). *On Tolerance: A Defence Of Moral Independence*. New York: Continuum International Publishing Group.

Gandy, K., King, K., Hurle, P. S., Bustin, C. & Glazebrook, K. (2016). *Poverty and decision-making. How behavioural science can improve opportunity in the UK.* London: Behavioural Insights Ltd.

Geertz, C. (1973). Thick description: toward an interpretive theory of culture. In M. Martin & L.C. McIntyre (Eds.), *Readings in the Philosophy of Social Science* (pp. 213-232). Cambridge, MA: MIT Press.

Gellner, E. (1985). The Psychoanalytic Movement: The Cunning of Unreason. Evanston: Northwestern University Press.

Gemeente Schoon. (2015). Serious game zwerfafval. Retrieved from: http://kenniswijzerzwerfafval.nl/document/serious-game-zwerfafval.

Gemeente Schoon. (2014). Stappenplan voor gedragsverandering. Retrieved from: http://kenniswijzerzwerfafval.nl/document/stappenplan-voor-gedragsverandering.

Giddens, A. (1998). The Third Way: The Renewal of Social Democracy. Cambridge: Polity Press.

Gigerenzer, G. (2015). On the supposed evidence for libertarian paternalism. *Review of philosophy and psychology*, *6*(3): 361-383.

Gladwell, M. (2006). Blink: The power of thinking without thinking. New York: Little, Brown and Company

Glennerster, R., & Takavarasha, K. (2013). *Running Randomized Evaluations: A practical guide*. Princeton: Princeton University Press.

Goffman, E. (1959). *The Presentation of Self in Everyday Life*. New York: Doubleday.

Goodwin, T. (2012). Why we should reject 'nudge'. *Politics*, *32*(2): 85-92.

Gowan, T. (2010). *Hobos, hustlers, and backsliders: Homeless in San Francisco*. Minneapolis: University of Minnesota Press.

Graf, R. (2019). Nudging Before the Nudge? Behavioural Traffic Safety Regulation and the Rise of Behavioural Economics. In H. Strassheim & S. Beck (Eds.), *Handbook of behavioural change and public policy*. Cheltenham: Edward Elgar Publishing.

Grimmelikhuijsen, S., Jilke, S., Olsen, A. L. & Tummers, L. (2017). Behavioral public administration: Combining insights from public administration and psychology. *Public Administration Review*, *77*(1): 45-56.

Grüne-Yanoff, T. (2012). Old wine in new casks: libertarian paternalism still violates liberal principles. *Social Choice and Welfare*, *38*(4): 635-645.

Hallsworth, M., Chadborn, T., Sallis, A., Sanders, M., Berry, D., Greaves, F., Clements, L. & Davies, S. C. (2016). Provision of social norm feedback to high prescribers of antibiotics in general practice: a pragmatic national randomised controlled trial. *The Lancet, 387*(10029), 1743-1752.

Hallsworth, M., Egan, M., Rutter, J., & McCrae, J. (2018). *Behavioural Government*. London, UK

Hallsworth, M., Parker, S. & Rutter, J. (2011). Policy-making In The Real World. London: Institute for Government.

Hallsworth, M., Snijders, V., Burd, H., Prestt, J., Judah, G., Huf, S. & Halpern, D. (2016). *Applying behavioral insights: Simple ways to improve health outcomes.* Dohar, Qatar: World Innovation Summit for Health.

Halpern, D. (2015). *Inside the Nudge Unit*. London: Ebury Publishing.

Halpern, D., Bates, C., Mulgan, G. & Aldridge, S. with Beales, G. & Heathfield, A. (2004). Personal responsibility and changing behaviour. London: Prime Minister's Strategy Unit

Halpern, D. & Gallagher, R. (2015). Can 'nudging' change behaviour? Using behavioural insights to improve program redesign. In J. Wanna, H. A. Lee & S. Yates (Eds.), *Managing Under Austerity, Delivering Under Pressure.* Canberra: ANU Press.

Hammersley, M. & Atkinson, P. (1983). *Ethnography: Principles in practice.* London: Tavistock.

Han, B. C. (2017). *Psychopolitics: Neoliberalism and new technologies of power*. London/Brooklyn: Verso Books.

Hansen, P. (2018). What are we forgetting? *Behavioural Public Policy, 2*(2): 190-197.

Hansen, P. G. & Jespersen, A. M. (2013). Nudge and the manipulation of choice. *The European Journal of Risk Regulation*, 1: 3-28.

Hausman, D. M. & Welch, B. (2010). Debate: To nudge or not to nudge. *Journal of Political Philosophy*, *18*(1): 123-136.

Haynes, L., Goldacre, B., & Torgerson, D. (2012). *Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials*. London: Cabinet Office Behavioural Insights Team.

Head, B. W. (2008). Three lenses of evidence-based policy. *Australian Journal of Public Administration*, *67*(1): 1-11.

Heclo, H. & Wildavsky, A. (1974). *The Private Government of Public Money*. Berkeley: University of California Press.

Hiscox, P. M., Oliver, T., Ridgway, M., Holzinger, L. A., Warren, A., & Willis, A. (2017). *Going Blind to See More Clearly: The Effects of De-Identifying Job Applications in the Australian Public Service. BETA Working Paper 2017-1*. Canberra: Commonwealth of Australia.

Hood, C. (1991). A public management for all seasons? *Public Administration*, *69*(1): 3-19.

Hood, C., & Peters, G. (2004). The middle aging of new public management: into the age of paradox? *Journal of Public Administration Research and Theory*, *14*(3): 267-282.

Hommes, S., Franssen, R., Dirven, L, Mastop, J. & Schyns, P. (2016). Klimaatbestendige tuinen en daken stap Doorgronden-Sanity check. Deltares. Retrieved from: https://www.deltares.nl/nl/nieuws/klimaatbestendig-tuinieren-en-groene-daken/

Hoppe, R. (2010). From 'Knowledge Use' Towards 'Boundary Work'. In R. in 't Veld (Ed.), *Knowledge Democracy. Consequences for Science, Politics and Media*. Berlin, Heidelberg: Springer.

House of Lords Science and Technology Select Committee. (2011). *Behaviour Change. Science and Technology Select Committee. 2nd report of session 2010-12.* London: Authority of the House of Lords.

Howlett, M., Perl, A. & Ramesh, M. (2009). *Studying Public Policy: Policy Cycles & Policy Subsystems*. Oxford: Oxford University Press.

Ingold, J., & Monaghan, M. (2016). Evidence translation: An exploration of policy makers' use of evidence. *Policy & Politics, 44*(2): 171-190.

Jacobs, J. A. & Frickel, S. (2009). Interdisciplinarity: A critical assessment. *Annual Review of Sociology*, *35*: 43-65.

Janis, I. L. & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment*. New York: Free Press.

Jankowski, P. J. (2007). Embrace: A theologically informed peace psychology. *Christian Scholar's Review 2007, 36*(2): 141-157.

Janssen, M., Bergsma, E. & Cooper, K. (2017). Positieve gedragsbeïnvloeding om inwoners te activeren. Eigen schuld… of niet? *Sociaal Bestek,* 5: 37-39.

Jellema S. R., Zantinge, E. M., Van der Lucht, F. (2014). *Nudge fuzz: different nudge definitions and an investigation of nudges promoting healthy behavior.* National Institute for Public Health and the Environment/Utrecht University.

John, P. (2014). Policy entrepreneurship in UK central government: The behavioural insights team and the use of randomized controlled trials. *Public Policy and Administration, 29*(3): 257-267.

John, P. (2017). Behavioural science, randomized evaluations and the transformation of public policy: The case of the UK government. In J. Pykett, R. Jones & M. Whitehead (Eds.), *Psychological governance and public policy: Governing the mind, brain and behaviour.* Abingdon: Routledge.

John, P. (2019). The international appeal of behavioural public policy: is nudge an Anglo-American phenomenon? *Journal of Chinese Governance*: 1-18.

John, P. (2018). *How far to nudge?* Cheltenham: Edward Elgar Publishing.

John, P., Smith, G. & Stoker, G. (2009). Nudge nudge, think think: two strategies for changing civic behaviour. *The Political Quarterly*, *80*(3): 350-361.

Johnson, B., & Hagström, B. (2005). The translation perspective as an alternative to the policy diffusion paradigm: The case of the Swedish methadone maintenance treatment. *Journal of Social Policy, 34*(3): 365-388.

Jones, K. S. (2017). The 'cognitive bias' of behavioural economics and neuropolitics. Retrieved from: https:// kittysjones.wordpress.com/2017/10/17/behaviour-economics-and-neuropolitics/

Jones, R., J. Pykett, & Whitehead, M. (2013). *Changing Behaviours: On the Rise of the Psychological State*. Cheltenham/Northampton: Edward Elgar Publishing.

Jones, R. & Whitehead, M. (2018). Politics done like Science: Critical Perspectives on Psychological Governance and the Experimental. *Environment and Planning D: Society and Space, 36*(2): 313-330.

Jonkers, P. & Tiemeijer, W. (2014). *Policymaking Using Behavioural Expertise*. Synopsis of the WRR report 92. The Hague: The Netherlands Scientific Council for Government Policy (WRR).

Jupp, E., Pykett, J. & Smith, F. M. (2016). *Emotional States: Sites and Spaces of Affective Governance*. London and New York: Taylor & Francis.

Kahneman, D. (2011). *Thinking, Fast and Slow.* London: Penguin Books.

Kahneman, D., Slavic, P. & Tversky, A. (Eds.) (1982) *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Kahneman, D. & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39*(4): 341-350.

Karré, P., Schillemans, T., van der Steen, M. & van der Wal, Z. (2017). *Toekomst van de bestuurskunde.* Den Haag: Boom Bestuurskunde.

Kaufman, H. (1981). *The Administrative Behavior of Federal Bureau Chiefs.* Washington: The Brookings Institution.

Kaufman, H. (1960). *The Forest Ranger.* Baltimore: Johns Hopkins Press.

Keltner, D., Oatley, K. & Jenkins, J.M. (2014). *Understanding Emotions*. Hoboken, NJ: Wiley.

KIM (Kennisinstituut voor Mobiliteitsbeleid). (2017.) *Stedelijke distributie en gedrag.* Den Haag: Ministerie van Infrastructuur en Milieu.

Kingdon, J. (1984). *Agendas, Alternatives, and Public Policies*. Boston: Little, Brown.

Kirchler, E. & Hölzl, E. (2006). Twenty-five years of the journal of Economic Psychology (1981-2005): A report on the development of an interdisciplinary field of research. *Journal of Economic Psychology, 27*: 793-804.

Kislov, R., Hodgson, D. & Boaden, R. (2016) Professionals as knowledge brokers: The limits of authority in healthcare collaboration. *Public Administration, 9*(2): 472-489.

Kleinpaste, T. (20 July 2013). Stiekem burgers manipuleren: van urinoir tot orgaandonatie. Publieke waarden zijn veranderd in driftbeïnvloeding. NRC.

Knight, C., & Lyall, C. (2013). Knowledge Brokers: The Role of Intermediaries in Producing Research Impact. *Evidence and Policy, 9*(3): 309-316.

Knorr-Cetina, K. (2007). Culture in global knowledge societies: Knowledge cultures and epistemic cultures. *Interdisciplinary Science Reviews*, *32*(4), 361-375.

Kooyman, M. (21 August 2018). Asofietsers Erasmusbrug buigen voor boegeroep en pijlen op de weg. Algemeen Dagblad. Retrieved from: https://www.ad.nl/rotterdam/asofietsers-erasmusbrug-buigen-voor-boegeroep-en-pijlen-op-de-weg~a1e416b0/

Kosters, M. & Van der Heijden, J. (2015). From mechanism to virtue: Evaluating Nudge theory. *Evaluation*, *21*(3): 276-291.

Kroese, F. (2017).*'Living lab: Nudging op het globe college'.* Research report.

Kruiter, A. J. (2010). *Mild despotisme: de hedendaagse democratie en verzorgingsstaat door de ogen van Alexis de Tocqueville*. Alkmaar: Van Gennep.

Lake, A. & Townshend, T. (2006). Obesogenic environments: exploring the built and food environments. *The Journal of the Royal society for the Promotion of Health*, *126*(6): 262-267.

Lamme, V. (2016). *Waarom? Op zoek naar wat ons werkelijk drijft.* Zaandam: Mind and Brain.

Lascoumes, P. & Le Galès, P. (2007). Introduction: understanding public policy through its instruments – from the nature of instruments to the sociology of public policy instrumentation. *Governance, 20*(1): 1-21.

Larson, M. S. (1977). *The Rise of Professionalism. A Sociological Analysis*. Berkeley: University of California Press.

Latour, B. (1986). The powers of association. In J. Law (Ed.), *Power, Action and Belief: A New Sociology of Knowledge?* London: Routledge.

Latour, B. (1987). *Science in Action: How to follow scientists and engineers through society*. Cambridge, Massachusetts: Harvard University Press.

Leggett, W. (2014). The Politics of Behaviour Change: Nudge, Neoliberalism and the State. *Policy and Politics 42*(1): 3-19.

Lepenies, R., Mackay, K. & Quigley, M. (2018). Three challenges for behavioural science and policy: the empirical, the normative and the political. *Behavioural Public Policy*. doi: 10.1017/bpp.2018.18

Lepenies, R. & Małecka, M. (2015). The institutional consequences of nudging - nudges, politics, and the law. *Review of Philosophy and Psychology, 6*(3): 427-437.

Lepenies, R. & Malecka, M. (2016). Nudges, Recht und Politik: Institutionelle Implikationen. *Zeitschrift Für Praktische Philosophie*, *3*(1): 487-530.

Lepenies, R. & Małecka, M. (2018). The Ethics of Behavioural Public Policy. In A. Lever & A. Poama (Eds.), *The Routledge Handbook of Ethics and Public Policy* (pp. 513-525). Abingdon: Routledge.

Lewis, O. (1968). The Culture of Poverty. In D.P. Moynihan (Ed.), *On Understanding Poverty: Perspectives from the Social Science*. New York: Basic Books.

Lin, Y., Osman, M., Harris, A. J. L. & Read, D. (2018). Underlying wishes and nudged choices. *Journal of Experimental Psychology: Applied, 24*(4): 459-475.

Lindblom, C. E. (1959). The Science of 'Muddling Through. *Public Administration Review*, *19*(2): 79-88.

Lindblom, C. E. (1990). *Inquiry and change.* New Haven, Connecticut: Yale University Press.

Lindblom, C. E. & Cohen, D. (1979). *Usable knowledge. Social science and social problem solving.* New Haven: 1979.

Lodge, M., & Wegrich, K. (2016) The Rationality Paradox of Nudge: Rational Tools of Government in a World of Bounded Rationality. *Law & Policy, 38*(3): 250-267.

Loewenstein, G., John, L. K. & Volpp, K. (2012). Using Decision Errors to Help People Help Themselves. In E. Shafir (Ed.), *The Behavioral Foundations of Public Policy* (pp. 361-379). Princeton: Princeton University Press.

Lourenço, J. S., Ciriolo, E., Almeida, S. R. & Troussard, X. (2016). *Behavioural Insights Applied to Policy*: *European Report 2016*. Publications Office of the European Union. Retrieved from: https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/behavioural-insights-applied-policy-european-report-2016

Lugosi, P. (2006). Between Overt and Covert Research: Concealment and Disclosure in an Ethnographic Study of Commercial Hospitality. *Qualitative Inquiry, 12*(3): 541-561.

Lunn, P. (2012). Behavioural economics and policymaking: Learning from the early adopters. *The Economic and Social Review*, *43*(3): 423-449.

Lunn, P. (2014). *Regulatory Policy and Behavioural Economics*. Paris: OECD Publishing.

Malecka, M. & Lepenies, R. (2018). Is the Behavioral Approach a Form of Scientific Imperialism? An Analysis of Law and Policy. In U. Mäki, A. Walsh & M. Pinto (Eds.), *Scientific Imperialism: Exploring the Boundaries of Interdisciplinarity*. Abingdon/New York: Routledge.

MacLeod, M. (2018). What makes interdisciplinarity difficult? Some consequences of domain specificity in interdisciplinary practice. *Synthese, 195:* 697-720.

MacLeod, M. & Nagatsu, M. (2018). What does interdisciplinarity look like in practice: Mapping interdisciplinarity and its limits in the environmental sciences. *Studies in History and Philosophy of Science Part A, 67:* 74-84.

Madden, R. (2017). *Being ethnographic. A guide to the theory and practice of ethnography*. London: SAGE Publications.

Martin, S. J., Goldstein, N. & Cialdini, R. (2014). *The small big: Small changes that spark big influence*. London: Hachette UK.

McGoey, L. (2012). The logic of strategic ignorance. *The British Journal of Sociology, 63*(3): 553-576.

McLaughlin, K. (2016). *Empowerment: A critique.* London: Routledge

Mettler, S. (2011). *The submerged state: How invisible government policies undermine American democracy.* Chicago and London: The University of Chicago Press.

Meyer, M. (2010). The rise of the knowledge broker. *Science Communication, 32*(1): 118-127.

Ministry of Economic Affairs. (4 December 2014). Kabinetsreactie op adviesrapporten van Rli, RMO en WRR over de benutting van gedragswetenschappelijke kennis in beleid. Retrieved from: https://www.eerstekamer.nl/behandeling/20141204/brief_regering_kabinetsreactie_op/info

Ministry of Economic Affairs (2016). *Rapport gedragsbeïnvloeding bij bedrijven.* The Hague: Ministry of Economic Affairs. Retrieved from: https://www.rijksoverheid.nl/documenten/rapporten/2017/01/11/rapport-gedragsbeinvloeding-bij-bedrijven

Mintrom, M. (1997). Policy entrepreneurs and the diffusion of innovation. *American Journal of Political Science, 41*(3): 738-770.

Mols, F., Haslam, S.A., Jetten, J. & Steffens, N. K. (2015). Why a nudge is not enough: A social identity critique of governance by stealth. *European Journal of Political Research, 54*(1): 81-98.

Mulderrig, J. (2017a). Reframing obesity: A critical discourse analysis of the UK's first social marketing campaign. *Critical Policy Studies, 11*(4): 455-476.

Mulderrig, J. (2017b). The language of 'nudge' in health policy: Pre-empting working class obesity through 'biopedagogy'. *Critical Policy Studies*, advance online publication. doi: 10.1080/19460171.2017.1398672

Mulderrig, J. (2018). Multimodal strategies of emotional governance: a critical analysis of 'nudge' tactics in health policy. *Critical Discourse Studies*, *15*(1): 39-67.

Newman, J. (2016). Deconstructing the Debate over Evidence-Based Policy. *Critical Policy Studies, 11*(2): 211-226.

Nimishakavi, S. (10 October 2016). The long-lasting wages of neglect: Flint residents plagued again by water crisis. *Non-Profit Quarterly*. Retrieved from: https://nonprofitquarterly.org/2016/10/10/long-lasting-wages-neglect-flint-residents-plagued-water-crisis/

Norman, D. (2002). *Emotion & design: attractive things work better. Interactions, 9*(4): 36-42.

Novi Mores. (2017). Handreiking voorkomen van bijplaatsingen. Retrieved from http://www.kenniswijzerzwerfafval.nl/download_document/1028.

Noordegraaf, M. (2007). From 'pure' to 'hybrid' professionalism. *Administration & Society*, *39*(6): 761-785.

Noordegraaf, M., Van der Steen, M. & Van Twist, M. (2014). Fragmented or connective professionalism? *Public Administration*, *92*(1): 21-38.

Nutley, S. Morton, S. Jung, T. & Boaz, A. (2010). Evidence and policy in six European countries: diverse approaches and common challenges. *Evidence & Policy, 6*(2): 131-144.

Nutley, S., & Davies, H. T. (2000). Getting research into practice: making a reality of evidence-based practice: some lessons from the diffusion of innovations. *Public money and management, 20*(4): 35-42.

OECD. (2017). *Behavioural Insights and Public Policy Lessons from around the World*. Paris: OECD Publishing.

Oliver, A. (2015). Nudging, shoving, and budging: Behavioural economic-informed policy. *Public Administration, 93*(3): 700-714.

Oliver, A. (2013). From nudging to budging: using behavioural economics to inform public sector policy. *Journal of Social Policy*, *42*(4): 685-700.

Oliver, A. (2017). *The Origins of Behavioural Public Policy.* Cambridge: Cambridge University Press.

O'Reilly, K. (2008). *Key concepts in ethnography.* Thousand Oaks, California: Sage.

Ossewaarde, R. (2007). The New Social Contract and the Struggle for Sovereignty in the Netherlands. *Government and Opposition, 42*(4): 491-512.

Ossewaarde, R. (2010). Keuzevrijheid als Sturingsinstrument. *Bestuurskunde, 19*(2): 5-11.

Overman, S. (2016). *Great expectations of autonomous agencies*. PhD dissertation. Radbound University, Nijmegen: Edward Elgar.

Paardekoper, C., Van Genugten, M.L. & Wesseling, H. (2013). Eerste overheid in kruiend ijs-van bestuurlijke vernieuwing naar transformatie van de natiestaat? *Bestuurskunde,* 3: 45-52.

Parsons, W. (2002). From Muddling through to Muddling up – Evidence Based Policy Making and the Modernisation of British Government. *Public Policy and Administration, 17*(3): 43-60.

Pedwell, C. (2017). Habit and the Politics of Social Change: A comparison of nudge theory and pragmatist philosophy. *Body & Society, 23*(4): 59-94.

Poiesz, T. (1999). *Gedragsmanagement*. Utrecht: AnkhHermes.

Pollitt, C., & Hupe, P. (2011). Talking About Government. *Public Management Review, 13*(5): 641-658.

Pressman, J. L. & Wildavsky, A. B. (1974). *Implementation*. Oakland: University of California Press.

Prij, J. (2018). Met zelfregie overvraagt overheid burgers. *Christen Democratische Verkenningen,* 2*:* 44-50.

Pykett, J. (2018). Geography and neuroscience: Critical engagements with geography's "neural turn". *Transactions of the Institute of British Geographers, 43*(2): 154-169.

Pykett, J., Jones, R., Welsh, M. & Whitehead, M. (2014). The Art of Choosing and the Politics of Social Marketing. *Policy Studies, 35*(2): 97-114.

Rayner, G., & Lang, T. (2011). Is nudge an effective public health strategy to tackle obesity? No. *British Medical Journal, 342*: d2177.

Reisch, L. A., Sunstein, C. R., & Gwozdz, W. (2017). Beyond carrots and sticks: Europeans support health nudges. *Food Policy, 69*: 1-10.

Reid, L. & Ellsworth-Krebs, K. (2019). Nudge(ography) and practice theories: Contemporary sites of behavioural science and post-structuralist approaches in geography? *Progress in Human Geography, 43*(2): 295-313.

Rentea, S. (2016). From the katechontic to the eschatological in contemporary liberal biopolitics. In S. Prozorov & S. Rentea (Eds.), *The Routledge Handbook of Biopolitics*. Abingdon: Routledge.

Rhodes, R., P. 't Hart & Noordegraaf, M. (2007). *Observing Government Elites.* London: Palgrave Macmillan.

Rhodes, R. A. W. (2014). 'Genre blurring'and public administration: What can we learn from ethnography? *Australian Journal of Public Administration, 73*(3): 317-330.

Rip, A. (2000). In praise of speculation. In OECD, *Proceedings, Social Sciences for Knowledge and Decision Making* (pp. 95-103). Paris: OECD.

Rietkerk, M., Hommes, S., Mastop, J., Dirven, L. & Schyns, P. (2016). Klimaatbestendige tuinen – stap Doorgronden: gedragsanalyse. Deltares. Retrieved from: https://www.deltares.nl/nl/nieuws/klimaatbestendig-tuinieren-en-groene-daken/

Rogers, E. (1983). *Diffusion of Innovations*. New York: Free Press.

Rouw, R. (2011). Gevoel voor bewijs. Den Haag: Nederlandse School voor Openbaar Bestuur (NSOB).

Rowson, J. (2011). *Transforming behaviour change: Beyond nudge and neuromania.* London: The Royal Society of Arts (RSA).

Rutter, T. (23 July 2015). The rise of nudge – the unit helping politicians to fathom human behaviour. The Guardian. Retrieved from: https://www.theguardian.com/public-leaders-network/2015/jul/23/rise-nudge-unit-politicians-human-behaviour

Ryan, J. D. (2017). *To what extent have the policy recommendations of the Behavioural Insights Team been in accordance with nudge theory*? Master's thesis. University of Twente.

Sabel, C. & Zeitlin, J. (2012). Experimentalist governance. In D. Levi-Faur (Ed.), *The Oxford Handbook of Governance* (pp. 169-183). Oxford: Oxford University Press.

Sanders, M. Snijders, V. & Hallsworth. M. (2018). Behavioural science and policy: where are we now and where are we going? *Behavioural Public Policy, 2:* 144-167.

Sanderson, I. (2002). Evaluation, policy learning and evidence-based policy making. *Public administration*, *80*(1): 1-22.

SBST (Social and Behavioural Sciences Team). (2015). *Annual Report 2015*. United States Government. Retrieved from: https://sbst.gov/download/2015%20SBST%20Annual%20Report.pdf

Scott, J. C. (1998). Seeing like a state. New Haven: Yale University Press.

Schatz, E. (2013). *Political Ethnography: What Immersion Contributes to the Study of Power.* Chicago: University of Chicago Press.

Scherpenisse, J., Feitsma, J. N. P. & Schillemans, T. (2016). *Pluriform en passend. Het inbedden van gedragskennis in beleidsvorming.* Utrecht/Den Haag: Utrecht School of Governance/Netherlands School of Public Administration (NSOB).

Schillemans, T. (2017). Staat van de bestuurskunde: Samenvattend en persoonlijk slotakkoord. *Bestuurskunde*, *2017*(1): 88-97.

Schillemans, T. & de Vries, G. (Eds.) (2016a). Gedragskennis in bestuur en beleid. *Bestuurskunde, 25*(3).

Schillemans, T., & de Vries, G. (2016b). De homo psychologicus op het schip van staat: Gedragskennis in bestuur en beleid. Bestuurskunde, 25(3): 3-8.

Schillemans, T. (2012). *Mediatization of Public Services. How Organizations Adapt to the News Media*. Frankfurt: Peter Lang.

Schlag, P. (2010). Nudge, Choice Architecture, and Libertarian Paternalism. *Michigan Law Review, 108*(6): 913-925.

Schneider, A. & Ingram, H. (1990). Behavioral assumptions of policy tools. *The Journal of Politics, 52*(2): 510-529.

Schonewille, A. (2015). *What activation practitioners do*. Vrije Universiteit Amsterdam: Ridderprint.

Schön, D. (1983). *The Reflective Practitioner*. New York: Basic Books.

Schuppert, G. F. (2016). Nudging: Neither a Novelty, nor a Promising Lead – Unless in Context. In A. Kemmerer, C. Möllers, M. Steinbeis & G. Wagner (Eds.), *Choice Architecture in Democracies* (pp. 333-338). Baden-Baden: Nomos.

Schubert, C. (2017). Exploring the (Behavioural) Political Economy of Nudging. *Journal of Institutional Economics, 13*(3): 499-522.

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

Selinger, E. & Whyte, K. P. (2012). Nudging cannot solve complex policy problems. *European Journal of Risk Regulation, 3*(1): 26-31.

Selznick, P. (1957). *Leadership in administration.* Berkeley: University of California Press.

Sent, E. M. (2004). Behavioural Economics: How Psychology Made Its (Limited) Way Back into Economics. *History of Political Economy,* 4: 735-760.

Service, O., Hallsworth, M., Halpern, D., Algate, F., Gallagher, R., Nguyen, S., Rudd, S., Michael, S., *with* Pelenur, M., Gyani, A., Harper, H., Reinhard, J. & Kirkman, E. (2014). *EAST: Four Simple Ways to Apply Behavioural Insights.* London: The Behavioural Insights Team.

SETUP. (2015). *Urban nudging.* Project report. Retrieved from: http://docplayer.nl/25585019-Projectverslag-urban.html

Shafir, E. (Ed.). (2012). *The Behavioral Foundations Of Public Policy*. Princeton: Princeton University Press.

Shankar, M. & Foster, L. (2016). *Behavioural insights at the United Nations - Achieving agenda 2030. United Nations Development Programme.* Retrieved from: https://www.undp.org/content/undp/en/home/librarypage/development-impact/behavioural-insights-at-the-united-nations--achieving-agenda-203.html

Shipan, C.R. & Volden, C. (2012). Policy diffusion. Seven lessons for scholars and practitioners. *Public Administration Review*, 6: 788-796.

Shove, E. (2010). Beyond the ABC: climate change policy and theories of social change. *Environment and Planning A, 42*: 1273-1285.

Shove, E., Pantzar, M., & Watson, M. (2012). *The dynamics of social practice: Everyday life and how it changes.* London: Sage Publications.

Silva, M. L. E. (2015). Queer sex vignettes from a Brazilian favela: An ethnographic striptease. *Ethnography, 16*(2), 223-239.

Simon, H. A. (1947). *Administrative Behavior: A study of decision-making processes in administrative organization.* Oxford: Macmillan.

Simon, H.A. (1957). *Models of Man: Social and Rational*. London: John Wiley and Sons.

Simon, H.A. (1985). Human nature in politics: The dialogue of psychology with political science. *The American Political Science Review, 79*(2): 293-304.

Simons, A. & Voss, J. (2018). The concept of instrument constituencies: accounting for dynamics and practices of knowing governance. *Policy and Society, 37*(1): 14-35.

Sørensen, E. & Torfing, J. (2009). Making governance networks effective and democratic through metagovernance. *Public Administration, 87*(2): 234-258.

Smith, G., & May, D. (1980). The Artificial Debate between Rationalist and Incrementalist Models of Decision Making. *Policy and Politics, 8*(2): 147-161.

Star, S. L. & Griesemer, J. R. (1989). Institutional ecology, 'translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science, 19*(3): 387-420.

Stevens, A. (2011). Telling policy stories: an ethnographic study of the use of evidence in policy-making in the UK. *Journal of Social Policy, 40*(2): 237-255.

Stillman, S. (23 January 2017). Can Behavioural Science Help in Flint. New Yorker. Retrieved from: https://www.newyorker.com/magazine/2017/01/23/can-behavioral-science-help-in-flint

Stinesen, B. B. & Renes, R. J. (2014). *Overheidsbeleid gericht op gedragsverandering: meer dan een mooi streven?* Den Haag: Raad voor Maatschappelijke Ontwikkeling.

St. Paul, G. (2011). *The Tyranny of Utility.* Princeton: Princeton University Press.

Strassheim, H. (2017). Bringing the Political Back in: Reconstructing the Debate over Evidence-Based Policy. A Response to Newman. *Critical Policy Studies, 11*(2): 235-245.

Strassheim, H., Jung, A. & Korinek, R. (2015). Reframing Expertise: The Rise of Behavioural Insights and Interventions in Public Policy. In A. Bertoin Antal, M. Hutter & D. Stark (Eds.), *Moments of Valuation. Exploring Sites of Dissonance*. Oxford: Oxford University Press.

Strassheim, H. & Korinek, R. (2015). Behavioural governance in Europe. In J. Wilsdon & R. Doubleday (Eds.), *Future directions for scientific advice in Europe.* Cambridge: Centre for Science and Policy.

Strassheim, H. & Korinek, R. (2018). The Politics of Nudging. In J. Allmendinger & U. Noack (Eds.), *WZB Report 2018.* Berlin: WZB (Berlin Social Science Center).

Strategic Policy Making Team Cabinet Office. (1999). *Professional Policymaking in the Twenty-First Century*. London: Cabinet Office.

Strauss, K. (2009). Cognition, context, and multimethod approaches to economic decision making'. *Environment and Planning A, 41:* 302-317.

Stone, D. (2012). *Policy Paradox: The Art of Political Decision Making*. New York/London: W.W. Norton & Company.

Stone, D. (2017). Understanding the transfer of policy failure: bricolage, experimentalism and translation. *Policy & Politics, 45*(1): 55-70.

Sunstein, C. R. (2014). Nudging: a very short guide. *Journal of Consumer Policy, 37*(4): 583-588.

Sunstein, C. R. (2017a). *Human Agency and Behavioral Economics.* Cham: Springer International Publishing.

Sunstein, C. R. (2017b). Misconceptions About Nudges. SSRN: https://ssrn.com/abstract=3033101

Sunstein, C. R., Reisch, L. A. & Rauber, J. (2018). A worldwide consensus on nudging? Not quite, but almost. *Regulation & Governance, 12*(1): 3-22.

Sunstein, C. R. & Thaler, R. H. (2003). Libertarian paternalism is not an oxymoron. *The University of Chicago Law Review, 70*(4): 1159-1202.

Tannenbaum, D., Fox, C. R. & Rogers, T. (2017). On the misplaced politics of behavioural policy interventions. *Nature Human Behaviour,* 1: 0130.

Thaler, R. H. (2015). *Misbehaving: The Making of Behavioural Economics.* London: Penguin Books.

Thaler, R. H. & Sunstein, C. R. (2008). *Nudge: Improving Decisions about Health, Wealth and Happiness*, London: Penguin Books.

Thorstendahl, R. (1990). *Essential Properties, Strategic Aims and Historical Development: Three Approaches to Theories of Professionalism*. London: Sage.

Tiemeijer, W. L. (2011). *Hoe mensen keuzes maken: de psychologie van het beslissen.* Amsterdam: Amsterdam University Press.

Tiemeijer, W. L., Thomas, C. A. & Prast, H.M. (Eds.) (2009). *De menselijke beslisser. Over de psychologie van keuze en gedrag.* The Netherlands Scientific Council for Government Policy (WRR). Amsterdam: Amsterdam University Press.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science, 185*(4157): 1124-1131.

Timmer, D. A. (1975). *Oversocialization and overpsychologization: a comparison of sociologism, psychologism and existentialism on the dialectic of individual and society*. Doctoral dissertation. University of Northern Iowa.

Towfigh, E. & Traxler, C. (2016). Nudges polarize! In A. Kemmerer, C. Möllers, M. Steinbeis & G. Wagner (Eds.), *Choice Architecture in Democracies* (pp. 321-328). Baden-Baden: Nomos.

Trommel, W. A. (2009). *Gulzig bestuur.* Den Haag: Boom Bestuurskunde.

Trommel, W. A. (2018). *Veerkrachtig bestuur: voorbij neoliberale drift en populistische kramp.* Den Haag: Boom Bestuurskunde.

Van Aaken, A. (2015). Judge the Nudge: In Search of the Legal Limits of Paternalistic Nudging in the EU. In A. Alemanno & A. L. Sibony (Eds.), *Nudging and the Law. What Can EU Law Learn from Behavioural Sciences?* Oxford: Hart Publishing.

Van Bavel, R., Herrmann, B., Esposito, G. & Proestakis, A. (2013). *Applying behavioural sciences to EU policy-making.* Brussels: Joint Research Centre European Commission.

Van Bavel, R. & Dessart, F. J. (2018). *The case for qualitative methods in behavioural studies for EU policy-making.* Luxembourg: Publications Office of the European Union.

Van der Meulen, M. (2009). *Achter de schermen. Vakontwikkeling en professionalisering van publieke managers in de zorg en bij de politie*. Delft: Eburon.

Van der Wal, Z. (2017). *The 21st Century Public Manager.* London: Macmillan International Higher Education.

Van Hulst, M. (2008). Quite an Experience: Using Ethnography to Study Local Governance. *Critical Policy Analysis, 2*(2): 143-159.

Van Oorschot, K., Haverkamp, B., van der Steen, M. & van Twist, M. (2013). *Choice architecture*. Working paper. The Hague: Netherlands School of Public Administration (NSOB).

Van Putten, R. J. & Trommel, W. A. (2018). Towards Public and Critical Approaches of Public Administration. Paper presented at Annual Conference Netherlands Institute of Governance 2018.

Van Staveren, I., Drenth, B., Zuure, J., Korthagen, I. & van Poppel, L. (2014). *De verleiding weerstaan. Grenzen aan beïnvloeding van gedrag door de overheid.* Den Haag: Raad voor Maatschappelijke Ontwikkeling (RMO).

Vugts, S., van den Hoven, M., de Vet, E. & Verweij, M. (2018). How autonomy is understood in discussions on the ethics of nudging. *Behavioural Public Policy*: 1-16. doi: 10.1017/bpp.2018.5

Waldron, J. (9 October 2014). It's all for your own good. *The New York Review of Books*. Retrieved from: http://www.nybooks.com/articles/2014/10/09/cass-sunstein-its-all-your-own-good/

Wallace-Wells, B. (13 May 2010). Cass Sunstein Wants to Nudge Us. *The New York Times Magazine.* Retrieved from: https://www.nytimes.com/2010/05/16/magazine/16Sunstein-t.html

Wansink, B., Painter, J. E. & Lee, Y. K. (2006). The office candy dish: proximity's influence on estimated and actual consumption. *International Journal of Obesity, 30*(5): 871-875.

Ward, V., A. House, & Hamer, S. (2009). Knowledge Brokering: The Missing Link in the Evidence to Action Chain? *Evidence and Policy,* 5: 267-279.

Weiss, C. H. (1979). The Many Meanings of Research Utilization. *Public Administration Review*, *39*(5): 426-431.

Weiss, C. H. (1993). Where Politics and Evaluation Research Meet. *Evaluation Practice, 14*(1): 93-106.

White, M. D. (2013). *The manipulation of choice*. New York: Palgrave Macmillan US.

Whitehead, M., Jones, R., Howell, R., Lilley, R. & Pykett, J. (2014). *Nudging all over the world.* Economic and Social Research Council of the United Kingdom.

Whitehead, M., Jones, R., Lilley, R., Pykett, J. & Howell, R. (2017). *Neuroliberalism. Behavioural Government in the Twenty-First Century*. Abingdon/New York: Routledge.

Whitehead, M., Jones, R., Howell, R., Pykett, J. & Lilley, R. (2018). Neuroliberalism: Cognition, Context and the Geographical Bounding of Rationality. *Progress in Human Geography,* advance online publication. doi: 10.1177/0309132518777624

Wilensky, H. L. (1964). The professionalization of everyone? *American Journal of Sociology*, *70*(2): 135-158.

Wilson, T. D. & Juarez, L. P. (2015). Intuition is not evidence: Prescriptions for behavioural interventions. *Behavioural Science & Policy, 1*(1): 13-20.

WINK Website. (2019). Retrieved from: http://winknudgingandbeyond.com/

Wintour, P. (9 September 2010). David Cameron's 'nudge unit' aims to improve economic behaviour. The Guardian. Retrieved from: https://www.theguardian.com/society/2010/sep/09/cameron-nudge-unit-economic-behaviour

Wolfe, S. E. (2012). Water Cognition and Cognitive Affective Mapping: Identifying Priority Clusters within a Canadian Water Efficiency Community. *Water Resource Management, 26:* 2991-3004.

World Bank. (2015). *Mind, Society and Behavior.* Washington: The World Bank.

Wulff, H. (2002). Yo-yo fieldwork: Mobility and time in a multi-local study of dance in Ireland. *Anthropological Journal on European Cultures, 1*(11): 117-136.

WRR (The Netherlands Scientific Council for Government Policy) (2017). *Weten is nog geen doen, Een realistisch perspectief op redzaamheid.* The Hague: WRR.

Yanow, D. (2000). *Conducting Interpretive Policy Analysis.* Thousand Oaks, California: Sage Publications.

Yeung, K. (2012). Nudge as fudge. *The Modern Law Review, 75*(1): 122-148.

# Appendices

*Appendix I: Additional observations in support of Chapter 7*

This appendix belongs to Chapter 7 'Institutionalizing The Behavioural State'. It provides additional observations from the conducted ethnographic fieldwork on Dutch behaviour experts, following the structure of Table 7.2 to 7.6 in the chapter.

**Ad Table 7.2: Evidence of rationalism**

| *Theme* | *Example* | | |
|---|---|---|---|
| Rationalist policy cycle | BIT EZ, the behavioural team of the Ministry of Economic Affairs follows the following working method in its projects:<br>1. Problem Analysis<br>2. Context Analysis<br>3. Possible Interventions<br>4. Testing<br>(BIN NL 2017) | BIT IenW, the behavioural team of the Ministry of the Infrastructure and Water Management, follows a similar method:<br>1. Unravelling<br>2. Designing<br>3. Experimenting<br>4. Monitoring and Evaluating<br>(BIN NL 2017) | A behaviour expert at the Rijkswaterstaat agency, uses an equivalent work method:<br>1. Determine problematic behaviour<br>2. Determine desired behaviour<br>3. Analyse behaviour<br>4. Determine behavioural measures<br>5. Implementation<br>6. Test with pre- and post-tests and control situation<br>(Gemeente Schoon 2014) |
| | An interviewee at a regulatory agency expands on the adopted working method in the team: 'At the bottom of practically every project lies a regulation strategy. These strategies didn't really exist before the behaviour experts were there, but since then they do. Such a strategy has a fixed format, a nine-step-action-plan (…) in which the role of behaviour is central. So, step one and two is problem exploration and problem analysis. Step three is effect measurement. Step four is causality. Five is subject segmentation or target group analysis. Six is the mix of enforcement. And then you've got planning, implementation, and evaluation. It's quite a chain of steps, but I say this because it shows that behaviour has at once been given a much more central position in our regulation. Because now we're specifically looking at it. And we're performing analyses first, before we make decisions like "we're going to inspect this hospital, or fine that general practitioner" because this is what happened in the past. Then you would just receive a signal, and on the basis of that sign you would act, yes or no. And now the most important problems are all being formulated in advance, and the most important target groups are distilled.' | | |
| | An interviewee at the Netherlands Enterprise Agency talks about his behavioural insights approach. He argues that many of his peers, embracing the 'Nudge' hype, use those insights in an uninformed way, as they forget to follow to earlier stages of the rationalist policy cycle: 'And often it [nudge] is being used by the government as a new panacea: let's nudge. The project at Economic Affairs used to be called the nudging project. And that I think is really the wrong type of approach, totally wrong. What the right approach would be? The right approach, I think, is that you don't automatically decide on your instruments, but first decide | | |

| | | | |
|---|---|---|---|
| | on: what do you actually want to change? And what in the behaviour has to change? First you've got to look at the behaviour change challenge, the behaviour that needs to change (…) and on that basis you make a proper behavioural analysis.' | | |
| | An interviewee at the 'Team Behaviour Change' of the Dutch Tax and Customs Administration explains her team's research approach as follows. They start with a formulation of the behaviour they want to achieve. This desired behaviour must be very concrete and measurable. Then they perform a behavioural analysis: what are the factors that are of influence on the desired behaviour? They draw both from psychology and from practical experience. The next steps are the development of interventions and testing their effects in a fashion that is as controlled as possible. | | |
| Belief in readability and craftability of human conduct | The Netherlands Enterprise Agency, following the Behaviour Change Wheel model, adopts ´COM-B' as its behavioural model. In this model, the B stands for Behaviour, and COM for the three main determinants of behaviour:<br>• Capability:<br>• Opportunity<br>• Motivation<br>(BIN NL 2017) | The Netherlands Authority for Consumers and Markets has a similar model:<br>• Capabilities<br>• Motivations<br>• Opportunities<br>• Resistance<br>• Biases and heuristics<br>(BIN NL 2017) | The Campaign Strategy Instrument, developed by the Ministry of General Affairs, identifies nine behavioural determinants:<br>• Social environment<br>• Emotions and associations<br>• Physical environment<br>• Self-image<br>• Capability<br>• Habits and automatisms<br>• Attitude<br>• Intention<br>• Knowledge<br>(BIN NL 2017) |
| | An interviewee elaborates on the essence of his approach: pinpointing what does and what doesn't trigger people to behave in certain ways: 'You just have to look very specifically, case-by-case, at what triggers people to do or not do things. Sometimes, that happens consciously, and sometimes unconsciously. You just have to look at what… which triggers cause people to move. Or not move.' | | |
| Hegemony of RCTs | An interviewee emphasizes the importance of testing behavioural designs because of the complexity and unpredictability of human behaviour. 'In the end, when you talk with people for a very long time about the theme "using behavioural insights in policymaking", you'll always find that you won't know how people will behave in advance. So you'll always have to test that.' | | |
| | An interviewee depicts running experiments as a core pillar of his behavioural practice: 'What you mainly want with behaviour change is examining and demonstrating that it works. (…) And because we've got this increase in data, there are much more behaviours (…) that we can map out. So, then you can start running experiments and effect measurements with A/B testing, and then you're able to see: this intervention works better than that one.' | | |
| | A local behaviour expert expresses the ambition to work more 'broadly', 'systematic' and 'evidence-based'. For her, the use of behavioural insights links closely to the evidence-based policy ambition to experimentally assess 'what works': 'I hope that I can apply some of those behavioural insights in the interventions (…) and also to add a bit of research and evidence-based policy thinking in the sense of: what works and what doesn't? | | |

**Ad Table 7.3: Evidence of knowledge brokering**

| *Theme* | *Example* |
|---|---|
| Smoothing the circulation of knowledge | An interviewee states: 'I'm not a social psychologist, nor a behavioural economist. That's always the discussion that you get into. I'm not an expert in that domain, in terms of content, my say in the matter is limited.' And later in the interview: 'I really am an uber-networker, I'm not an expert'. |
| | A behaviour expert at the 'Team Behaviour Change' of the Dutch Tax and Customs Administration notes that strong networks are crucial to get things done: '[T]here are quite some things that we don't think are working so well and that's always difficult, because at the Tax and Customs Administration a great deal of parties are responsible for one single thing. (…) It always tends to be somewhat of a struggle to make your point from a psychological perspective or to achieve something. You've got to have just the right people together. The Tax and Customs Administration very much is a network organization, so if you've got a good network then it's easier to roll out things on a large scale than when you don't. So you really need to bring together the right people.' |
| | An interviewee at a regulatory agency depicts his job as forming as a bridge between behavioural science and the daily practice within his field. He helps his peers with translating abstract knowledge into practical and contextualized solutions. 'I think it's an illusion to think that you have generic knowledge with generic solutions with these kind of things [behavioural insights]. You notice with these topics that most of all it's important to pay attention to understanding how the whole thing works [behavioural analysis]. (…) I think that's the most important part in which we function as a bridge. We've got methods for that.' |
| Facilitating concrete application | An interviewee at a regulatory agency depicts his job as forming as a bridge between behavioural science and the daily practice within his field. He helps his peers with translating abstract knowledge into practical and contextualized solutions. 'I think it's an illusion to think that you have generic knowledge with generic solutions with these kind of things [behavioural insights]. You notice with these topics that most of all it's important to pay attention to understanding how the whole thing works [behavioural analysis]. (…) I think that's the most important part in which we function as a bridge. We've got methods for that.' |
| Politico-strategic work | In 2017, the City of Enschede collaborated with a renowned Dutch behaviour change consultancy bureau in the area of financial debt. It applied several nudging techniques in a letter towards debt holders and ran an experiment to test their effects. The new letter resulted in significantly more debt holders undertaking action (i.e. paying, partially paying, arranging a settlement, or getting in touch) to pay their debts. After running the trial, a group of involved local officials reported about the project results in a Dutch professional journal (Janssen et al. 2017). While this fulfils an informational function, there may also be an implicit political dimension to it, explicitly showcasing successes in order to gain wider support and secure future resources. The following quote of an interviewee who was involved in the project hints at this political dimension of having to 'sell' one's approach, and the importance of clear-cut results in doing so. 'Of course we've had the luck that the project has become a big success, and we gladly share this success with others: Look, a 30% higher yield.' In selling the approach, the project members constructed a clear-cut 'what works' narrative, putting the spotlights on a core finding – the '30%' – that best demonstrated the project's effectiveness. This finding was for instance centrally presented in the opening paragraph of the article: 'In Enschede we experimented with a new collection letter according to the principles of "nudging". The |

| | |
|---|---|
| | result: the chance that a debt holder undertakes action increases with thirty percent.' (Janssen et al. 2017: 37). A subtlety about this result is that it is doesn't specifically explain the effect on the nudge-letter on the amounts of debt holders who paid back their debts. Instead, it explains more widely how many extra debt holders undertake *any sort of relevant action* to pay back their debts, which also includes, for instance, getting in contact with the City. An extra, overarching research variable of 'undertaking action' was constructed for this in the analysis. By constructing this more generic variable, and using it in the research analysis, it was possible to 'find' a relationship with a bigger effect size than the effect size that would have been found by using the other singular variables. Other than this, it ought to be noted that this 30% effect was a *relative* increase, whereas the *absolute* increase in effect size between the control group and treatment group in the field experiment was 11% (Janssen et al. 2017). While indeed this absolute effect size was *also* noted in the article, it is the relative effect size that is mentioned far more often and presented as the core finding. These are clear signs of efforts to accentuate the findings that most strongly frame the experiment as successful. |
| | Interestingly, some behaviour experts use their self-propagated behavioural insights as a vehicle for their internal political operations. They perform analyses on why policymakers do (or do not) engage with behavioural insights, seeking to extract ways to get them engaged with behavioural insights. In one organization, a trainee was hired to further explore how policymakers could be engaged. A behaviour expert elaborates on this internally-oriented 'meta-use' of behavioural insights as follows: 'We're trying to change this organization, and the behaviour of employees within this organization. (…) We've also looked from a behavioural perspective at: how can we nudge this organization, so that it will start doing this [use behavioural science]? |

**Ad Table 7.4: Evidence of output focus**

| *Example* |
|---|
| Two interviewees at a regulatory agency elaborate on their team's successes in terms of having come up with better regulation strategies.<br>Interviewee 1: 'In a general sense, people are very positive, although the troublesome thing is that you can't measure that very well…'<br>Me: 'What's your own feeling about it?<br>Interviewee 1: I think, what I said before, [our success lies in] the shift from only sanctioning at the end to also preventing at the start.<br>Me: 'So, the success then is having a better strategy?'<br>Interviewee 1: 'Yes, in more thorough regulation.<br>Interviewee 2: 'Yes, but also in thinking better instead of directly doing things. Actually, we're forcing our colleagues to think about… what do you want to achieve with this [their suggested design]? And why do you think that this is the accurate instrument to achieve that? And not this other instrument?' |

**Ad Table 7.5: Evidence of analytical satisficing**

| *Example* |
|---|
| Many behaviour experts speak about their analytical findings in a language of 'semi-certainty', emphasizing the informed but simultaneously inconclusive nature of their findings. They talk about *'educated guesses', 'indications', 'rules of thumb',* and findings that are not 'wholly pure' or 'statistically sound'. The following regulatory behaviour expert for instance expresses some caution regarding the conclusiveness of the data that is gathered: 'And in that way you're to some extent able to map out what could be possible characteristics of possibly violating groups. These are still not rock-hard data, but they are the data that we build upon when we give shape to our regulatory strategy.' |
| A regulatory behaviour expert questions the feasibility of organizing his field more according to a scientific logic, more strongly grounded in the accumulation of 'hard' knowledge. In his eyes, this leads to an unproductive 'effect measurement stress', aspiring to accumulate a level of scientific certainty that is rather impossible in a field that is bounded in its time, capacities, and resources. A looser, more experimental approach, working with softer knowledges is deemed more appropriate: 'Yes, yes… then you arrive at the question: how "hard" is it all? And however much I'd want to say that everything should be hard, it's just not feasible to find all things out in a scientific hard manner. (…) These things aren't engineerable, you know. It's about trial-and-error. So in that sense, the question is whether you should try to aim for hard truth-finding or whether you should organize your processes so that work more by trial-and-error. So, much more short policy cycles of evaluating and adapting.' |
| Two interviewees describe how they make *'educated guesses'* in their behavioural analyses. Despite the degree of uncertainty in those guesses, they still perceive such analyses as adding to the effectiveness of policies.<br>Me: 'Isn't it highly difficult to make that educated guess about whether there, for instance, is inertia going on or whether it's about following the status quo?<br>Interviewee 1: It's not simple but I think that with all the information we've got, we often come a long way. And yes, it remains an educated guess.<br>Interviewee 2: It's difficult to test whether it's accurate, that part…<br>Interviewee 1: You'll only know if it has an effect after the intervention, when you run a measurement, insofar as that's possible. It gives you information for future cases.' |

**Ad Table 7.6: Evidence of the horizontalization of the hierarchy of evidence**

| *Theme* | *Example* |
|---|---|
| Deemphasizing the value of RCTs | In the eyes of this regulatory behaviour expert, RCTs may be preferable but are less feasible in practice. His field is not an 'ideal testing environment': 'We can never really afford ourselves the luxury to experiment with different treatment groups, to explore what makes some people behave differently and others not. (…) The only thing we can often do is to beforehand try to make a guess as accurately as possible - well, it's a little bit more than a guess – about what are the determinants of the behaviour and how we can tap into those effectively? And of course we'll look into what the effects are, to draw lessons for the next time. But it's almost all of the time terribly difficult, if not impossible, to say: "Why don't we go and run some different interventions to see what happens?"' As this interviewee explains, part of why RCTs are infeasible is his inability to set up a controlled experiment; if he experiments on one treatment group of agencies, the others are likely to find out, which will contaminate the research design. Moreover, his organization not unusually performs its regulatory activities under the radar, meaning that he cannot openly run interventions on targeted organizations or involve them in field experiments. |

| | |
|---|---|
| | When prompted about the use of RCTs, a regulatory behaviour expert portrays RCTs as preferable yet also downplays their feasibility. This lack of feasibility is related to information gaps and dynamic field environments. 'I can imagine very well that it's very dependent on the information position that you've got as an agency. That's something that we're really noticing. Look, ideally you'll do everything with pre- and post-tests, but then you've got to possess the relevant data. And when you go measure something, for instance an intervention, you've got to [be able to assess] the effectiveness thereof. Then it has to concern a purely isolated intervention in a relatively stable environment. And if you look at our field for instance… well, those type of interventions don't or only rarely take place. That makes it difficult.' |
| | A behaviour expert problematizes the use of experiments from a juridical perspective, claiming that the randomization occurring with experiment designs goes against the juridical imperative to treat policy subjects equally: 'If we're really going to start regulating and enforcing, there simply can be no question of arbitrariness. (…) So then you're unable to work with a control group, an experimental group, or even multiple [treatment] groups. Because that would mean that you are operating arbitrarily, as you're selecting that group randomly whereas you're regulating, and that's simply not allowed. So it's not so much an ethical dilemma; it's simply prescribed by the law that it's not allowed.' |
| | This behaviour expert deemphasizes the usability of RCTs due to a mix of concerns. He notices that the experimental practice of creating different treatment groups, targeting some organizations while leaving others unmonitored, can be ethically undesirable. This, in combination with transferability issues, feeds the gap between the worlds of academic science and regulatory practice: 'Look, ideally you'll need natural experiments, but you can't always regulate half of the market and let the other half untouched to see what happens. That's not possible. And that's simply what restrains most of the testing of interventions. (…) [Talking about experimentation in his specific field] [I]t's also the case that it's so stylized that the question is whether the findings extend to the field in the first place. RCT's are therefore just one tool in the kit of the researcher, next to comparative studies and time-series analysis etcetera.' |
| | This behaviour expert thinks RCTs are sometimes preferable yet also rather costly and technically difficult procedures providing a degree of rigor that is not always needed. Interviewee: 'Well, that [running RCTs] is not really my thing, but it's necessary of course if you want to prove certain things… that it works… and to which extent it works. But it's difficult and costly to set up such an experiment in the, let's say, cruel-and-dirty outer world. Where for example neighbours interact with each other, making the randomization impure. That's what we're struggling with now as well. (…)'<br>Me: 'Is that the reason that you say it's not your thing… because it's just technically difficult…?'<br>Interviewee: 'Well yea, exactly, and often it's about getting the precise effect and the quantification thereof. And then I think… there has to be a significant effect… and you can also observe that in different ways than through (…) very pricy and time-consuming trials. And that's often possible. [For instance] when [referring to an ongoing project about waste collection] those people during their collection activities notice something like: "Hey, this works, this really makes a big difference." Or: "We're collecting so and so much more." So you're going back to [relying on] estimations and key indicators… you can say something about those fairly quickly. I understand that in some cases you've really got to go to the RCT but that won't always be the case I think.' |

| | |
|---|---|
| Reemphasizing the value of non-experimental knowledge | A local behaviour expert notes that while in-depth analysis and experimentation is generally part of his repertoire, this is not always the case. When he gives ad hoc advice to his peers about their outgoing letters, he operates much more loosely informed, based on previous experiences and generic insights rather than thorough analysis and trials. 'Of course you'll always apply nudges in an experimental way, because you never know if it… [works] (…) Normally when you want to nudge, one of the essential things is that you make a proper analysis of the target audience. Obviously, you don't do that with those letters. You involve the target audience, but that's where it kind of stops. You try to adapt those letters with the knowledge that you've already acquired.' |
| | According to an interviewee, coming up with behavioural designs is not a mere matter of scientific analysis. It also requires a step of creative thinking, appealing to a source of knowledge that is less typically acknowledged in behavioural policy: 'Does the right, sound analysis automatically correspond with an appropriate intervention? No… that actually has been kind of my misunderstanding. I wasn't so aware of this (…) but it's what I assumed. It's what lots of officials assume. We're all academics, and we've all come to grasp reality in such a manner. Most officials are social scientists. Their way of reasoning and achieving solutions is to start with an analysis, collect all facts and knowledge about an issue; find the appropriate science about behaviour; construct a theory about the cause of the problem and about the relevant actors. That should almost automatically point to solutions, to interventions in which the causes of a problem in relevant actors are targeted. But this is not how it works, most certainly not when we talk about using behavioural insights in policy. That really requires a bit more of creative thinking. My eyes have really kind of been opened about that there is no linear logic from cause to solution. You don't get there just through making sound analyses. Which is not to say analyses is superfluous – it's very necessary. But not enough.' |
| | A local behaviour expert explains how his embrace of behavioural science differs from a scientific approach, and is more based on his own practical intuition. When asked how he knows whether his interventions actually work, he responds: 'You don't. That's intuition. And I think that's where it stops for us. I really mean that. I think that's where it ends for us. (…) How great would it be if we'd know that the colour 'red' triggers people's motivation to take the stairs, for instance? How great would that be? But in a different way, that's also a utopia and you'd have to experimentally assess that. You'd really have to do it: the experiment. And on a small-scale that may be possible but on the large scale it's mostly still about intuition though. (….) Maybe it's the case that… you could do some literature study (…) but it remains a cost-benefit matter. (…) There are many other things that are very nice and interesting. And the benefit for me already is a change in my way-of-thinking. And that's enough.' |

*Appendix II: List of focus group participants*

On 28 June 2018, at the Netherlands School of Public Administration in The Hague, a focus group was conducted to validate overarching findings of this study. The following behaviour experts participating during this focus group:

| *Name* | *Affiliation* |
|---|---|
| Chris Verhoeven | City of Utrecht |
| Thomas Dirkmaat | Ministry of Economic Affairs and Climate Policy (EZK) |
| Whitney Tanihatu | Dutch Healthcare Authority (NZa) |
| Ruth Gonschorrek | The Netherlands Authority for Consumers and Markets (ACM) |
| Kaj Bots | Employee Insurance Agency (UWV) |

# Samenvatting in het Nederlands

Een zogeheten 'Behavioural Insights' beweging heeft haar intrede gedaan in de mondiale beleidswereld. Deze beweging introduceert inzichten uit de *new behavioural economics school*, die aantonen dat mensen slechts beperkt rationeel zijn. Dit impliceert dat hun gedrag gestuurd kan worden door in te spelen op onbewust plaatsvindende beslisprocessen ('nudging'). Ook wijst het op het belang van de doorlopende evaluatie van de effecten van beleidsinterventies middels rigoureuze veldexperimenten. De toegevoegde waarde van Behavioural Insights is tot nog toe het onderwerp geweest van een raadselachtig debat. Pleitbezorgers zien in Behavioural Insights de belofte van effectiever, rigoureuzer en minder dwingend beleid. Critici daarentegen zien het als manipulatief, technocratisch en methodologisch beperkt. Ondertussen zetten sceptici het weg als een noch innovatieve noch impactvolle beweging. Het blijft daarmee onduidelijk wat Behavioural Insights nu daadwerkelijk te bieden heeft voor beleid en maatschappij.

Het ontbreekt echter vaak aan empirische diepgang in het debat over en onderzoek naar de opkomende 'gedragsoverheid'. Slechts het topje van de ijsberg lijkt zichtbaar. De aandacht gaat vaak uit naar het idee en ideaalbeeld van nudg*es,* niet zozeer naar de concrete bezigheden van nudg*ers* – terwijl die ons juist iets kunnen leren over hoe gedragsinzichten daadwerkelijk toegepast worden in de praktijk. Deze studie werpt daarom het licht op praktijken van pionierende gedragsexperts. Een nog maar weinig verkend terrein van Behavioural Insights wordt onder de loep genomen: het Nederlandse gedragslandschap. Waar bevinden de Nederlandse gedragsexperts zich, wie zijn ze en wat doen ze daadwerkelijk? Hoe verloopt hun professionalisering? Hoe 'kijken' ze? Deze vragen zijn verkend aan de hand van etnografisch veldwerk gedurende een periode van vier jaar, bestaande uit onder meer interviews, participant-observaties en een detachering bij een Behavioural Insights Team.

Deze studie laat zien dat de Nederlandse gedragsoverheid op het moment een bescheiden aanwezigheid vormt. Een groot deel van de dagelijkse praktijk bestaat uit het 'makelen' ofwel 'aan de man brengen' van kennis, in tegenstelling tot het direct toepassen ervan. Het veld is nog sterk in ontwikkeling en haar professionele grenzen zijn nog onbepaald. Het gedragslandschap is bovendien rijk, vol spanning en ambiguïteit. Waar gedragsexperts doorgaans een krachtig *frontstage* verhaal vertellen van rigoureus en gedragsbewust beleid, is de praktijk achter de schermen dikwijls meer pragmatisch en gefragmenteerd. *Backstage* is het aan gedragsexperts om de belofte van gedragsbewust beleid zo goed als mogelijk waar te maken in een vaak warrig, gehaast, onzeker en gepolitiseerd beleidsproces. De geobserveerde rijkheid in het gedragslandschap roept de vraag op hoe deze beweging zich verder zal ontwikkelen. Deze studie schetst en evalueert drie mogelijke toekomstscenario's – 'daad bij het woord', 'het wilde westen' en 'gecontinueerde hypocrisie' – elk gepaard met unieke baten en risico's. De mogelijkheid van een 'getemde' ontwikkeling wordt verkend, waarin de gedragsoverheid haar belofte van gedragsbewuster beleid kan waarmaken zonder door te schieten in een overpsychologisering van beleid.

# List of publications

*Publications linked to this dissertation*

**Feitsma, J. N. P.** (2018). The behavioural state. Critical observations on technocracy and psychocracy. *Policy Sciences, 51*(3): 387-410. doi: 10.1007/s11077-018-9325-5

**Feitsma, J. N. P.** (2018). 'Rationalized incrementalism'. How behavior experts in government negotiate institutional logics. *Critical Policy Studies,* advance online publication. doi: 10.1080/19460171.2018.1557067

**Feitsma, J. N. P.** (2019). Brokering behaviour change: the work of behavioural insights experts in government. *Policy & Politics, 47*(1): 37-56. doi: 10.1332/030557318X15174915040678

**Feitsma, J. N. P.** (2019). Meters maken: gedragsexperts in het lokaal bestuur. *Bestuurswetenschappen, 73*(1): 5-29. doi: 10.5553/Bw/016571942019073001002

**Feitsma, J. N. P.** & Schillemans, T. (2019). Behaviour experts in government: from newcomers to professionals? In H. Strassheim & S. Beck (Eds.), *Handbook of Behavioural Change and Public Policy.* Edward Elgar Publishing.

Ball, S. & **Feitsma, J. N. P.** (2019). The boundaries of Behavioural Insights: observations from two ethnographic studies. Revised and resubmitted.

**Feitsma, J. N. P.** & Whitehead, M. (2019). Bounded interdisciplinarity: critical interdisciplinary perspectives on context and evidence in behavioural public policies. Revised and resubmitted.

*Other publications*

**Feitsma, J. N. P.** & Schillemans, T. (2015). Gelukspolitiek en de wetenschappelijke uitdaging. *Liberaal Reveil,* 56(2): 102-105.

**Feitsma, J. N. P.** (2016). Meer dan een nudge. Gedragsexperts bij de Nederlandse overheid. *Bestuurskunde,* 2015, 3: 4-32. doi: 10.5553/Bk/092733872016025003004

Scherpenisse, J., **Feitsma, J. N. P.**, Van Twist, M., Schillemans, T. (2016). *Pluriform en passend - Het inbedden van gedragskennis in beleid.* NSOB/USG. Report for the Ministry of the Interior and Kingdom Relations.

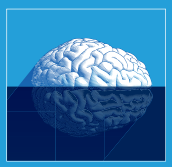
**Feitsma, J. N. P.** (2018). De makelaarsmachinerie achter gedragsbeleid. *Christen Democratische Verkenningen, 2*: 74-79.

# About the author

Joram Feitsma was born in 1990 in Amsterdam and grew up in Grou, a small village in Friesland, in the northern part of the Netherlands. He started his academic career in 2008 studying Public Administration and Organization Science at Utrecht University School of Governance (USG), combined with the two-year interdisciplinary Descartes College honours program at Utrecht University. He holds two master's degrees, one in Strategic Human Resource Management, and one in Philosophy, both received from Utrecht University. As part of the Philosophy master, Joram did coursework and research at Washington University, St Louis.

Subsequently, Joram began his doctoral research based at USG, conducting an ethnographic study on the emerging 'Behavioural Insights' movement in governments. He temporarily worked as policy adviser for a Dutch ministerial Behavioural Insight Team as part of the fieldwork. He has published in *Critical Policy Studies, Policy Sciences, Policy & Politics*, and the *Edward Elgar Handbook of Behavioral Change and Public Policy.* He has been a visiting researcher at the University of Queensland as well as Aberystwyth University, where he is academic partner of the Interdisciplinary Behavioural Insights Research Centre.

Next to his doctoral research, Joram has had various teaching, training and super/intervision responsibilities at USG, including the supervision of several master's students in writing their thesis. Outside of academia, he is an avid soccer player and composes and performs neoclassical piano music.



A 'Behavioural Insights' movement has entered the global policy scene. Drawing from the behavioural sciences, this movement has been pushing novel forms of policy analysis and design, particularly promoting the use of 'nudges' and field experiments. Despite Behavioural Insights' apparent popularity, its contribution to state and society has been appraised divergently. Some see a promise of radical evidence-based government, others a looming manipulative technocracy, and again others a trivial fad. In light of the puzzle posed by these clashing appraisals, this study has explored Behavioural Insights from up close. Based on ethnographic fieldwork in Dutch government, it asks what Behavioural Insights experts actually do, how they are professionalizing and how they 'see'. The result is a rich account of the emerging 'behavioural state', with a sharper understanding of its varieties, tensions and ambiguities.

Joram Feitsma is a qualitative researcher based at Utrecht University School of Governance. Trained in public administration and philosophy, he studies the 'behavioural turn' of governments.