



Power and Narratives in the Era of Behavioral Government

Manuel C. Ortiz de Landázuri*, Alejandro Cid Moreno**

* Associate Professor, Department of Philosophy, University of Navarra, Pamplona, Spain. E-mail: mcruz@unav.es

** Research Assistant, Marketing and Business Ethics, IESE University of Navarra, Barcelona, Spain. E-mail: icidmoreno@alumni.unav.es

Abstract

Behavioral government employs tools like nudges to steer behavior toward certain practices, raising ethical concerns if the underlying narratives are not transparent or inclusive. In this context, it has been debated whether these types of policies are a source of manipulation or rather are practices inherent to a libertarian paternalism. This article attempts to answer this question from an analysis of power. In contrast to theorizations of power in terms of force or pure domination, it is possible to establish a theoretical framework that understands power in an inclusive manner, as a capacity for order, and from there establish a rational framework for evaluating the narratives implicit in this type of policies. We propose that, within a framework of dialogue and shared values, power need not be coercive to be effective. The discussion concludes that narratives, when aligned with the common good and grounded in clear reasoning, legitimize behavioral government as an ethical and collaborative approach to addressing contemporary social challenges.

Keywords

choice architecture, libertarian paternalism, manipulation, nudges

1 Introduction

In recent years, there has been a growing interest in behavioral aspects that affect not only economics but also politics and government. There are already many institutions within governments that deal with behavioral policies, such as the Behavioural Insights Team (BIT) in the UK, the Behavioural Economics Team of the Australian Government (BETA) in Australia, the Office of Behavior and Public Policy (OCPP) in Spain, or the behavioral policy unit of the World Bank. And, in line with these initiatives, by “behavioral government” we are referring to decision-making based on unconscious and not entirely rational tendencies in terms of efficiency. It is for this reason that some authors have emphasized in recent decades the role played by unconscious aspects in decision-making (Ariely, 2008) and public opinion (Sunstein, 2019).

Narratives can be defined as structured stories and representations that give shape to a series of events, ideas, or experiences in a meaningful way. They serve as a framework to organize information and communicate meaning. In this article, we want to show the importance of narratives in the exercise of power in the light of what, in recent years, has been called “behavioral

government”. To do so, we first give some examples of narratives present in policies based on behavioral government. In this way, we show that behind these measures always underlie a narrative that serves for the constitution of power in a double dynamic: on the one hand, behavioral government is based on a previously admitted narrative, from which information is processed. On the other hand, behavioral governance generates a new narrative through new policies (reframing). The question arises, then, whether behavioral governance decisions hide a manipulative intention or only top-down guidance.

In this article, we address this question through the analysis of power and try to delimit what we mean by manipulation. Power has often been equated with the use of violence or domination, but from other perspectives, it is possible to understand power as a source of order, and thus establish parameters that ensure shared communicative processes and narratives based on good reasons. The risk of behavioral governance lies in the elaboration of narratives that appeal only to unconscious emotions and ignore other important aspects, such as coherence and shared values.

A particularly controversial aspect in this respect is whether behind the practices of behavioral governance lies a form of unconscious manipulation and domination. The central tension in behavioral governance lies in its dual nature: it can simultaneously serve as a tool for the collective good and as a mechanism for manipulation. Moving beyond the binary debate of whether ‘nudging’ is inherently manipulative or paternalistic, the ethical distinction can be found in the role of narratives as a cognitive and political bridge. Employing a Dual Process Theory framework (Kahneman, 2011), it becomes evident that while choice architecture primarily targets System 1 (intuitive processing), the presence of a transparent and shared narrative allows for System 2 (deliberative processing) engagement. This prevents the exercise of power from becoming a ‘clandestine’ influence and transforms it into a legitimate ‘power-to’ – a collective capacity where citizens can consciously endorse policy goals.

Thus, narratives are not mere rhetorical additions but the very condition that ensures the agency of the subject within the era of behavioral government. In this article, we explore the role of the narrative in establishing a behavioral framework that affects the decision-making of both rulers and ruled in order to discuss the basis of a soft “libertarian paternalism”. To this end, we first of all show, through some examples, in what behavioral governance consists, and reflect on the extent to which it is supported by certain narratives. Secondly, we analyze what constitutes power and what role narratives play in its exercise. Finally, we establish a theoretical proposal of the exercise of power through narratives in relation to the mechanisms implicit in behavioral governance. In the light of the narrative paradigm, it can be observed that power is developed through narrative mediations in which both the rulers and the ruled are involved.

And we are considering power in behavioral governance, because it is the most subtle and efficient form of soft power. This type of power is exercised through implicit narratives and could be seen as a form of manipulation. Literature on power has often emphasized coercive power (“power over” others), overlooking power as the capacity to achieve certain ends, or “power to” (Battezzorre, 2017). Power, thus understood, is an intrinsically negative concept, signifying that “A affects B in a manner contrary to B’s interests” (Lukes, 1974, 27). Structural power, from this perspective, manifests as a set of predominant values, beliefs, rituals, and institutional procedures that systematically benefit certain individuals and groups at the expense of others (Bachrach & Baratz, 1970, 43).

We find this equation of power and domination over others in many authors, sometimes understanding that domination is an act of getting someone to do something he or she would not want to do (Dahl, 1957), at other times understanding that, in the act of domination, there

is a force that leads to acting against one's own will (Wartenberg, 1990, 85). For example, according to Pettit (1996, 578), "one agent dominates another if and only if he or she has a certain power over that other: in particular, the power to interfere in the affairs of the other and to inflict a certain damage". As we shall see at the end of the article, this view of domination as the imposition of harm on another is the one that has permeated much of the discussion on the manipulation that behavioral governance can entail.

However, these analyses of power sometimes overlook that power is often exercised as follows: "A gets B to do something that is good for him and for the community." Power as dominance has usually been understood as A being able to impose on B something that he or she would not really want to do, but this presupposes that we are individuals who have interests prior to any form of political community. Many times, both A and B participate in certain shared narratives, and A's power lies in B's participation in a good, common to both (Parsons, 1967, 308). In this sense, power-to (power as capacity) is more important than power over others, because power over others makes sense only if it helps to achieve some ends. From Arendt's (1998, 200) perspective of power, it can be understood that not only individuals but also communities are subjects of power to the extent that there are shared narratives. Thus, behavioral governance need not be seen as manipulation, as long as clear intercommunicative processes are established. As such, Fisher's (1989) narrative paradigm may thus prove useful in establishing criteria for a good story: background values, internal coherence, and being based on good reasons.

In this article, we will first present a vision of power that goes beyond the framework of domination in order to understand how narratives allow for the construction of non-repressive and non-manipulative forms of power. Next, we will analyze specific characteristics of behavioral governance and nudges. Finally, we will examine how these types of practices can be understood through certain underlying narratives, and by what criteria they can be said to constitute manipulation or not.

2 Power and Intermediations

First, to comprehend behavioral governance, one must understand that it is a form of power and, as such, is constituted through certain narratives. However, power remains a diffuse and difficult concept to define. Here, we argue that power should be understood in general terms as *dynamis* or the capacity to accomplish things, and within human communities, as the capacity to order and shape actions toward an objective. In this sense, although power is linked to domination, the latter need not be repressive or violent; rather, it can be based on intermediations and shared narratives that integrate different agents around a common project, rooted in certain accepted values.

Literature on power has typically focused on its coercive aspects or understood domination in terms of violence and imposition (Lukes, 1974, 27; Bachrach & Baratz, 1970, 43; Wartenberg, 1990, 85; Pettit, 1996, 578). In this respect, even a significant figure like Foucault (1980), who clearly established the relationship between power and discourse, often frames power in terms of domination.

Due to Foucault's considerable influence on the understanding of power, we will briefly analyze the link he establishes between power and discourse before highlighting some inadequacies of his coercive model. Following this, and drawing on Han's (2005) critique of the purely coercive model of power, we advocate for a vision aligned with the perspectives

of Parsons and Arendt: power as the capacity to integrate forces around a shared objective or narrative. Power is the capacity to achieve some ends, and it grows in as much there is a shared narrative. Thus, it will be shown that behavioral government can be understood as a practice that is not necessarily manipulative.

As for Foucault, he focused a good part of his studies on the relationship between truth and power, and affirmed that power should be analyzed in its expressions, and one should not try to define it, since one would then be trapped in the nets of the modern science that he is criticizing (Foucault, 1997, 13–14). Foucault holds that power and truth are two linked concepts: power builds truth; truth legitimizes power. In this binary structure, power-knowledge truth is nothing more than a construct. Through his analysis of the constitution of sciences, he tries to show that truth is the result of a persuasive discourse that generates control. In the background, there is only power in terms of control and domination (Foucault, 1997, 15–30). And the role of reason in the constitution of power is to generate truth in terms of persuasion.

According to Foucault (1980, 109–116; 2014, 113), the modern state operates as a means of controlling the social body through norms and legal structures. The ruler wields power through violence and coercion to impose their will, with sovereignty shifting from control over territory to control over a population, akin to managing a natural body (Foucault, 2009, 23). This view of power has been very influential in the way we approach power relations, not only within society but also in international relations. However, although Foucault understands power as creative energy, the coercive model weighs heavily in his model.

In contrast to Foucault, we have in recent years the position of Byung-Chul Han (2005), who argues that real power has more to do with persuasion than with disciplinary methods of control and must be understood as the capacity to expand one's will to another's will. Han (2005) states that, although power has traditionally been interpreted through a coercive-negative model, greater power lies in influencing another's future rather than hindering it. Power does not consist solely in the capacity to use violence or impose punishments. Although that capacity may seem important to ensure order, it is equally true that respect for the law arises from the recognition of a juridical and social order that deserves to be upheld. Therefore, "the one who is only able to set his decision with the use of force or a negative sanction has little power" (Han, 2005, 25). Power configures networks and opens fields between people, in which different individual powers are organized toward a determined direction (Han, 2005). The essence of power, as Han observes, is not violence or domination, but the energy that helps to develop the political community. In this sense, there is no real power without communication and intermediations (Arendt, 1998).

On the other hand, as some scholars have argued (Porter, 2009; Sassen, 2006), it might be more accurate to say that power is not exercised by a single actor but is better understood through "assemblages" when analyzing governance. There are no independent and autonomous actors; rather, communities are governed through a network of institutions and practices in which actors move and play their roles. "Originally used in art, an assemblage is created out of disparate elements, each of which has meanings or purposes that might be quite unrelated to the other elements, but which together are brought into a new relationship with one another to create an arrangement with its own distinctive meaning or purpose" (Porter, 2009, p. 89). In this framework, there is no single sovereign power but a social epistemic network that generates forms of power through norms, practices, and institutions (Adler & Bernstein, 2005).

Governance takes place through a variety of interconnected networks. Rather than conceptualizing different institutions as sovereign powers, one should think in terms of networks of power involving states, non-state actors, and various intermediaries. These networks operate through systems of inclusion and exclusion procedures (Castells, 2011) and entrench their power

through communicative processes. For example, a standard defines the specific way in which a group of people is interconnected within a network. It acts as the shared norm or practice that enables network members to access one another (Grewal, 2008).

Power can be described as creative energy, or a “capacity to do”, which is actualized through networks and intermediations, with narratives playing a critical role in this process. From the perspective of Fisher’s (1989) narrative paradigm, humans primarily comprehend through stories rather than strictly logical discourses. These narratives incorporate imaginative elements and, more importantly, values and “good reasons” that resonate with individuals. Fisher identifies several key characteristics of the narrative paradigm, emphasizing that human communication is inherently narrative. He argues that, throughout history, rationality has often been reduced to logic, neglecting other forms of understanding, such as storytelling. Within stories, values are communicated through the narration of facts, and narrative rationality can be assessed based on coherence, fidelity, implicit values, and compelling reasons.

From this perspective, power is most effective when its narrative is well-constructed: when it reflects shared values and provides compelling, fact-based reasons, as narratives act as tools of unity and persuasion, enabling collective action. As Arendt (1998, 244) points out, power arises from shared narratives expressed through actions and words. Therefore, control or coercion without dialogue lacks the legitimacy of real power.

These narratives are shaped by various institutions and processes, reflecting broader social, cultural, and political contexts. In recent years, global narratives have gained prominence, influencing individual decision-making and policy development. Examples include ecological sustainability, gender equality, and bio-sanitary protection, which form part of global epistemic communities (Haas, 1992). These narratives establish codes and norms that structure power relations on a global scale, operating through networks rather than traditional hierarchies (Castells, 2011).

The thesis we defend here is that the narrative paradigm allows us to understand the exercise of power in the behavioral era not necessarily as a form of manipulation and despotic domination, but as the ordered orientation through shared narratives.

3 What is Behavioral Governance?

Traditional conceptions of governance often prioritize the explicit legislative and fiscal deliberations of political actors. However, contemporary regulatory practice suggests that individuals are increasingly governed not through formal prohibitions, but through subtle interventions in their decision-making environments. By configuring the “choice architecture”, ranging from the strategic placement of urban infrastructure to the framing of information in public spaces, governments exert a form of influence that bypasses conscious deliberation, systematically steering behavioral outcomes without resorting to coercive legal frameworks.

Behavioral governance concerns the processes of information assimilation and decision making so that various actors and population segments act in a certain way (Gofen et al., 2021). These are public policies aimed at generating an architecture of choice (Thaler & Sunstein, 2008). This type of governance assumes rationality from above, insofar as it is assumed that agents do not possess complete information when making decisions and need a little nudge or framework for action to make correct choices.

As some scholars have shown (Battaglio et al., 2019; Moseley & Thomann, 2021; Simon, 1978), individuals often tend to make decisions that are not the best precisely because of deficiencies in

the process of assimilating and interpreting information. We often act not on the basis of reality itself, but because of an inadequate perception of reality (Rittel & Webber, 1973; Thomann et al., 2018). Behavioral governance, then, can be understood as the set of psychological processes aimed at affecting the behavior of individuals (see Galizzi, 2017; Grimmelikhuijsen et al., 2017). In this framework, nudges, understood as small nudges or *stimuli* that unconsciously guide the activity of individuals, have gained special interest (Thaler & Sunstein, 2008; Sunstein, 2014).

The policy narrative, as a discursive tool, not only shapes public perception of policies but also profoundly influences the behavior of citizens and policymakers. In this framework, the concepts of “nudging” and “choice architecture”, analyzed by Lehner et al. (2015), and the conception of the state as a “market maker”, proposed by De la Cruz Prego (2021), highlight the intersection between policy design, human behavior, and sustainability. Lehner et al. (2015) present nudging as a behavioral approach to overcome cognitive limitations and human biases, promoting socially desirable decisions without coercion. These strategies have been successfully applied in domains such as energy efficiency and food consumption, demonstrating that tools such as social norms or changes in default choice are highly effective in changing specific behaviors. However, their effectiveness varies by context, underscoring the importance of culturally and socially tailored design.

In this line of thought, De la Cruz Prego (2021) argues that the state must assume an active role as an “architect of markets” to achieve the Sustainable Development Goals (SDGs). This approach implies not only regulating existing markets but also redesigning them to promote environmental sustainability and social equity objectives. The integration of behavioral tools, such as those proposed by Lehner et al. (2015), complements traditional interventions by incorporating knowledge of human biases and behaviors in the design of public policies. Both approaches converge on the need for policy narratives that legitimize these interventions, since a policy narrative not only explains and justifies policies to the citizens but also influences how policymakers perceive their own rationale and decision-making. This emphasizes the importance of integrating behavioral tools with robust institutional design and narratives that reinforce public trust in the policy process.

In the next section, we are going to show some examples of policies based on behavioral governance to demonstrate two things: on the one hand, there is always a narrative underlying these policies to shape the architecture of choice; on the other hand, it is not only the governed who are affected by these narratives, but also those who design public policies.

4 Behavioral Governance and Public Policy

Political narratives impact the formulation and reception of public policies, not only through large-scale decision-making, but sometimes through mechanisms that generate different scenarios and guide the behavior of individuals in certain directions. Let us look at some examples, then.

The automatic enrollment in pension schemes in the UK serves as a quintessential case study. This resulted in a significant increase in participation from 61% to 88% in four years, adding more than 5 million new savers (Halpern & Sanders, 2016). This shift was based on the design of default policies that use inertia to encourage financially responsible decisions. The narrative behind this policy framed the lack of savings as a “failure” of individual behavior, justifying state intervention as a tool to correct this bias. When analyzed through the narrative-power framework, this policy does not merely function as a mechanical default setting; rather,

it relies on a latent narrative of “long-term security” and “intergenerational responsibility”. While critics might view this as a bypass of individual will, this critique must account for the fact that non-enrolment equally constitutes a choice architecture, one that carries its own implicit narrative, namely, that saving for retirement is an individual responsibility requiring active initiative. The relevant distinction, therefore, is not between a “neutral” default and an “interventionist” one, but between narratives that operate transparently and those that remain unexamined.

Both enrolment and non-enrolment frames shape behavior; what differs is whether the underlying values are made visible and open to public scrutiny. The success of the nudge is contingent upon its alignment with a shared social narrative that legitimizes the state’s role in securing future welfare. From a behavioral symmetry perspective, this narrative omits that the experts designing these policies may also be influenced by optimism or overconfidence biases. While the policy is effective in terms of outcomes, it perpetuates an institutional reliance on individual decisions, limiting the space for citizen autonomy.

Madrian and Shea (2001) have analyzed how automatic enrollment in 401(k) deferred compensation plans affects savings behavior in a large U.S. corporation. Prior to implementing this policy, employees had to actively enroll to participate, resulting in a participation rate of 37%. After the introduction of automatic enrollment, this figure increased to 86%, even with no change in benefits or plan terms. In addition, 61% of automatically enrolled employees maintained the default contribution (3%) and funding allocation (100% in the money market fund), evidencing strong behavioral inertia. These results highlight how defaults not only simplify decisions but also act as implicit suggestions perceived as recommendations. It is a clear example of the power of biases, such as change aversion and procrastination, in policy design (for an overview of such behavioral tools in public policy, see Chetty, 2015). However, it raises ethical questions about the use of these mechanisms, as they shape decisions by exploiting cognitive limitations rather than encouraging more informed participation. As we will see below, the question is whether or not the unconscious orientation of individuals toward certain decisions is manipulation.

Similarly, interventions designed to increase fiscal compliance through social proof are often interpreted solely as psychological triggers. In Hallsworth et al.’s (2017) field experiments with the UK tax authority, messages telling recipients that “nine out of ten people pay their tax on time”, and, more strongly, that they were currently in the small minority who had not yet paid, raised payment rates by around five percentage points and accelerated several million pounds in tax revenue (Hallsworth et al., 2017). However, within the conceptual framework of this article, these interventions are understood as attempts to reinforce a “narrative of fairness”. By highlighting peer behavior, the state is not merely manipulating a bias, but is instead invoking a collective narrative where tax compliance is a shared contribution to the common good. Under this interpretation, the efficacy of the nudge depends on whether the citizen perceives this narrative as a legitimate expression of “power-to” (collaborative capacity) rather than an arbitrary imposition.

Another clear example of behavioral policy can be found in the nutritional traffic light, which has been promoted by many governments (although not in a prescriptive manner) and is having a major impact on food consumption (Osman & Thornton, 2019).

However, the narrative underlying behavioral government policies is not always the right one. Sometimes policies have been implemented based on erroneous assumptions, as for example in the case of predicting problematic alcohol consumption (Miller et al., 2024). Contrary to what was thought, delay discounting (the decrease in the value of a reward as its receipt is delayed) is not as significant a factor as the expectation of receiving rewards under temporal uncertainty,

or, in other words, human beings usually act more out of fear of losing than out of expectation of gaining.

One controversial example can be found in the letters that governments send to women over 50 to prevent breast cancer through mammography (with a specific appointment already in the letter). Although the letter states that mammography reduces mortality by 20%, this is actually an exaggerated figure (Gigerenzer, 2014). What this measure aims to do is to increase the rate of mammography, but not to teach people. The narrative underpinning policies based on these findings emphasizes the need to “correct” impulsive behaviors through interventions designed to improve decision-making. However, from a behavioral symmetry perspective, these policies could amplify expert design biases, particularly if paternalistic approaches are prioritized over participatory approaches.

As discussed above, political narratives not only shape citizen perception but also the rationality of policymakers. These narratives, combined with cognitive biases such as confirmation or authority bias, generate a cycle in which narratives and policy decisions reinforce each other. This is seen in the examples mentioned above: the use of social norms in tax charters has increased tax compliance by 5% (Hallsworth et al., 2017), or policies such as automatic enrollment in 401(k) plans have leveraged inertia and change aversion to improve citizens’ savings (Madrian & Shea, 2001). Behind these behavioral policies, one can read a narrative in which it is understood that the state must ensure that citizens contribute to the common good through taxes, energy savings, food health, and concern for the ecosystem. These narratives, where good and bad appear, are not generated directly by those who take measures, but are developed through communication networks in which many agents and processes intervene (Castells, 2011).

5 Behavioral Governance: Manipulation or the Need for Order?

Thaler and Sunstein (2003) have been the main proponents of nudging and behavioral government, coining the concept of “libertarian paternalism”, a view that grants freedom to individuals while assigning the state a role in channeling individual wills toward the common good. This perspective arises from the recognition that individuals often lack sufficient information to make optimal decisions. Since all policies inevitably influence individual choices, the state must at least nudge individuals through small motivations to achieve outcomes beneficial to both themselves and society, such as contributing to the common good, improving health, or achieving energy savings. This approach seeks to guide decisions toward positive outcomes while respecting freedom of choice, acknowledging the biases inherent in both individuals and policy design.

In contrast, the Behavioral Symmetry Theory developed by Thomas (2019) critiques the assumption that policymakers are inherently rational and benevolent agents, unlike citizens, who are often portrayed as subject to cognitive biases. This theory challenges the presumption that experts are free from systematic errors, arguing that behavioral interventions amplify the risks of manipulation if policymakers’ own biases are not acknowledged. This critique aligns with the concept of the policy narrative, which structures the perception of problems and solutions for both citizens and policymakers. As the Behavioural Insights Team report emphasizes, these narratives not only justify specific interventions but also shape how governments prioritize problems and evaluate their effectiveness (Hallsworth et al., 2018). In sum, Thomas’s (2019) Behavioral Symmetry Theory posits that both policymakers and citizens are influenced by cognitive biases, thereby questioning the superior rationality often attributed to experts.

It is important to note, however, that Thomas's critique does not require policymakers to be perfectly rational in order for nudge policies to serve citizens' interests. The relevant question is not whether experts are free of cognitive bias, but whether their informed judgements improve outcomes relative to unassisted individual decision-making. Consulting domain experts, whether physicians, economists, or public health specialists, is a rational strategy precisely because expertise reduces, even if it cannot eliminate, systematic error. Moreover, democratic accountability mechanisms – parliamentary oversight, freedom of information, and independent policy evaluation – serve as institutional correctives when expert judgement fails. Acknowledging the cognitive limitations of policymakers strengthens, therefore, rather than undermines, the case for transparent and externally scrutinized behavioral governance.

Thus, we have come to a key question, namely, whether there is an objective behind these practices to guide while preserving individual freedom, or do they represent ideological manipulation? Libertarian paternalism, as proposed by Thaler and Sunstein (2003), acknowledges that both citizens and policymakers are subject to cognitive biases and conditioned by the architecture of decisions. It suggests designing policies that respect individual freedom while nudging people toward choices that benefit their personal and collective well-being. This framework accepts the inevitability of decision conditioning, advocating for public policies that guide decisions toward social welfare without unduly restricting autonomy. However, as White (2013) has argued, libertarian paternalism presupposes that individuals are irrational and must be guided, potentially overlooking alternative motivations that might drive their actions. This raises an important question: is behavioral governance an exercise in manipulation based on arbitrary narratives, or is it better understood as an exercise of structured power? To address this, we will examine the foundations of power and its relationship to narrative.

The thesis we will maintain here is that these practices are always the product of a narrative or help to construct a narrative, so they can never be neutral. However, the fact that there is a narrative does not mean that there is manipulation. A correct approach to power shows that having power is not the same as exercising violence. Social organization inherently requires mechanisms of authority and coordination. The challenge lies in ensuring that such authority does not devolve into manipulation or coercive dominance. This transition is prevented when power is embedded within a shared and transparent narrative, established through robust dialogic processes. While critics such as White argue that individual decision-making remains independent of unconsciously assimilated frameworks, it is argued here that agency always operates within a narrative perspective – one often internalized without explicit awareness. This cognitive reality applies equally to policy designers, who operate from within settled institutional narratives that precede their own agency.

To understand how narratives transform behavioral governance from mere manipulation into a legitimate exercise of power, it is essential to ground the analysis in Dual Process Theory (Kahneman, 2011). Standard behavioral interventions primarily target System 1 – the fast, instinctive, and emotional brain – by exploiting cognitive biases to bypass deliberation. Critics argue that this “by-passing” constitutes manipulation because it excludes the individual's agency. However, narratives function as a bridge to System 2 – the slower, more deliberative and logical mind. While a nudge may provide the initial impulse (System 1), a transparent and shared narrative provides the reasons and values that allow System 2 to endorse the action. By providing a narrative framework, the government moves beyond “clandestine” influence and offers a justification that the citizen can rationally evaluate and consciously adopt as part of a collective “power-to” (capacity), rather than a coercive “power-over”.

At this point, we can see how behavioral governance does not necessarily have to be a form of manipulation. Although manipulation remains a somewhat diffuse concept (Wilkinson, 2013) due to the multiple methods through which it is carried out (Noogle, 2026), we could define manipulation as a behavior that seeks to direct someone's conduct toward the interest of the manipulator, concealing true intentions and nullifying the other person's capacity for free and informed choice. Whereas in persuasion both parties usually know the objective and can benefit, in manipulation, there is an instrumentalization of the person; they are turned into a means to an external end.

Contrary to those who believe that behavioral governance and, specifically, nudges are a form of manipulation (Grüne-Yanoff, 2012), it can be argued that they are instead a type of situational persuasion when they meet certain requirements. Specifically, behavioral governance practices can be subjected to a disclosure test (Dowding & Oprea, 2024). According to these authors, a communicative practice (or a nudge) is not manipulation if it meets two rules: (1) it offers a Suitable Reason: the influencer must believe that the reason offered is a good reason for the other person to act; and (2) there is Testimonial Honesty: the influencer must truly believe in the truth of what they are saying or presenting.

The hypothetical disclosure test poses a counterfactual question, namely: would the subject continue to act in the same way if the influencer openly revealed their intentions and the methods being used? If the subject proceeds, the influence is minimal or non-existent, because the subject accepts the reasons provided as valid for their own decision-making. If the subject changes their mind or feels indignant, it is because a high degree of manipulation exists, as the behavior depended entirely on concealment. For example, if a school cafeteria places fruit at eye level so that children eat better and, upon being told, "we have placed the fruit here so you see it first and eat healthily", the child says "oh, okay" and continues to choose the fruit, there is low manipulation. In contrast, if the child feels deceived or coerced upon discovering the trick, the degree of manipulation is greater.

Dowding and Oprea's (2024) proposal is especially relevant when analyzing how shared narratives legitimize behavioral governance, precisely because a narrative is "shared" and legitimate not because the people repeat it, but because it passes the Hypothetical Disclosure Test. If, upon revealing the narrative behind the nudge, the citizen recognizes it as their own and, based on acceptable reasons, then there is no manipulation, but rather a coordinated exercise of power.

From Kahneman's (2011) perspective, this represents an exploitation of System 1: the manipulator leverages our automatic responses and intuitive biases so that System 2 – our rational watchdog – does not even wake up.. Dowding and Oprea's (2024) test acts here as the ultimate intruder detector: if, upon turning on the lights of transparency and revealing the trick, the subject would abort the mission, we are facing pure manipulation.

Now, if this is the case, it is crucial for agents who design behavioral governance policies and practices to first consider which shared narratives their own practices are based on, as it is only from these shared narratives that different behavioral strategies can be viewed as persuasion rather than manipulation.

6 Narrative Paradigm and the Generation of an Unconscious Narrative

If the thesis of the narrative paradigm is correct, human beings do not primarily understand or make decisions based on logical arguments or rational discourses. Instead, we rely on narratives that we unconsciously assimilate, carrying implicit values with which we identify and which

shape our actions. The strength of behavioral governance lies in its ability to subtly connect specific individual actions to collective narratives, such as “together we all contribute to the common good”, “we must protect the planet”, or “taking care of your health benefits general welfare”. These narratives act as anchors that orient behavior without explicit coercion, fostering a sense of shared purpose and responsibility.

Consider issues like sustainability and ecology. Here, narratives construct a moral framework where citizens are cast as “heroes” taking active steps to save the planet, while those who disregard environmental concerns are subtly positioned as “villains” threatening collective survival. This dynamic drives behaviors such as recycling, reducing energy consumption, and supporting green policies. Similarly, in the domain of nutrition, narratives oppose industrial practices that prioritize profit over health and environmental sustainability. They encourage individuals to resist “easy consumption” of highly processed foods, advocating instead for balanced diets that align with ethical values and long-term well-being. These examples highlight how narratives create a framework for behavior, tying individual actions to larger societal goals through shared values and implicit moral codes.

Power, as exercised through behavioral governance, cannot exist without narrative. In this context, narratives are often more unconsciously assimilated than in other forms of governance. Importantly, the relationship between power and narrative in behavioral governance affects not only individuals but also policymakers, who are themselves influenced by the narratives they absorb. For instance, global narratives around climate change or public health often shape policy priorities, creating frameworks that guide decision-making at institutional levels. These frameworks are not neutral but embed specific values and visions of the common good, reinforcing the power of shared narratives in structuring both governance and public behavior.

For these reasons, behavioral governance should not necessarily be viewed as manipulative. Manipulation implies the use of coercion or dominance to achieve a benefit at someone else’s expense, often without transparency or consent. In contrast, the exercise of authority through behavioral governance can introduce an order based on good reasons, fostering behaviors that align with the common good. Human beings are interdependent and frequently lack the information or resources to guide behavior autonomously. In such cases, behavioral governance provides a framework for orientation, offering subtle guidance that respects individual agency while promoting collective well-being. However, this approach relies on ethical considerations, ensuring that narratives are transparent, inclusive, and directed toward the shared benefit of society as a whole.

Behavioral governance, when grounded in shared narratives and directed toward the common good, represents a tool that can align individual behavior with collective goals. It leverages the human tendency to find meaning through stories, creating a powerful mechanism for shaping societal dynamics while maintaining the delicate balance between guidance and autonomy.

As such, we claim that behavioral government is not necessarily a form of manipulation that removes freedom from citizens, because life itself demands that those who govern will have to make decisions for us and try to organize society in the best possible way. Power is violent when it goes against our own good, but there is a common good that demands that someone try to establish it through narratives. The risk of behavioral government is that it becomes a manipulative form of politics. In order to avoid excesses it is necessary to have certain mechanisms of transparency and control, such as external audits, but above all, it is necessary that the narratives underlying these practices are based on shared values and are open to dialogue.

7 Conclusions

Throughout this article, we have explored the theoretical foundations of the exercise of power in behavioral governance to highlight several key issues. First, power is always exercised through implicit narratives, and in the case of behavioral governance, it involves orienting behavior through unconscious mechanisms. These mechanisms work by connecting individuals to shared values and collective goals, embedding decisions within a broader narrative framework that subtly guides action.

Given that power operates through shared communicative networks and processes, the logic of behavioral governance need not be seen as manipulation, provided that the underlying narratives are oriented toward the common good and not merely toward achieving individual benefits. Power is not an inherent property of structures, as Dowding (2019, 9) wants us to believe, but is realized through the communities generated within these structures. These communities become subjects of power when shared values and narratives guide governance and decision-making processes. Narratives, therefore, serve as the foundation for collective action, fostering cohesion and legitimacy in governance practices.

The narrative paradigm emphasizes that there is logic and rationality embedded in discourses that do not conform to formal logical structures. Fisher (1989) argues that human understanding is inherently narrative, shaped by stories that carry implicit values and good reasons. These narratives are assimilated unconsciously and influence our decisions in profound ways. For instance, narratives about sustainability, health, or social responsibility shape consumer behavior, encouraging choices that align with broader societal goals. Far from being irrational, these decisions reveal a hidden logic rooted in shared narratives and collective values. Ultimately, the unconscious elements that drive behavior should not be dismissed as irrational but understood as part of a broader rationality that incorporates narrative-based reasoning. Behavioral governance, when grounded in ethical narratives oriented toward the common good, can serve as a powerful tool for fostering collaboration and social cohesion. By recognizing the role of narratives in shaping decisions, we can better understand how power operates within behavioral governance frameworks and ensure that it supports equitable and sustainable outcomes.

In conclusion, the ethical viability of behavioral government depends on its ability to engage both cognitive systems of the citizenry. While choice architecture inevitably influences System 1 through intuitive and automatic heuristics, it is the shared narrative that provides the necessary bridge to System 2 deliberation. By embedding nudges within transparent, inclusive narratives, the exercise of power shifts from a “power-over” (manipulation of the unconscious) to a “power-to” (collective capacity). A narrative that is publicly accessible and rationally defensible ensures that behavioral interventions are not mere techniques of subversion, but tools for agency that respect the individual’s capacity to consciously endorse the common good.

To move from theory to practice, behavioral governance must adopt institutional mechanisms that guarantee narrative openness. This includes (1) narrative audits, where independent bodies evaluate the ethical alignment of policy frameworks: a concept rooted in the tradition of frame-critical analysis in public policy (Schön & Rein, 1994). And (2) it must endorse a publicity principle, which mandates the disclosure of the underlying values of every nudge to invite System 2 scrutiny, grounded in the Kantian principle of publicity and developed within the Rawlsian concept of public reason (Rawls, 1993). Institutionalizing narrative audits ensures, on the one hand, that the justificatory framework of a policy is subject to external scrutiny, preventing the slide from persuasion into manipulative propaganda. On the other hand, a “publicity principle” mechanism requires that every behavioral intervention be accompanied by a clear statement

of its narrative goals, allowing System 2 engagement through public contestation (here the hypothetical disclosure test could serve as an effective tool for discerning when a practice aligns with a shared narrative, Dowding & Oprea, 2024). These mechanisms provide the “watertight” link between cognitive science and democratic legitimacy, ensuring that power remains a collaborative capacity.

Finding an appropriate form of behavioral governance represents one of the critical challenges of the twenty-first century. In an era where digital media and algorithmic systems play an increasingly influential role, governments must not only regulate these technologies but also provide guidance that aligns individual behavior with societal goals. This guidance, however, must avoid the pitfalls of manipulation. The key lies in ensuring that behavioral governance is exercised transparently, through participatory communication processes that foster shared narratives oriented toward the common good. By grounding governance in ethical principles and inclusivity, it is possible to harness the potential of behavioral governance without compromising individual autonomy or trust.

As societies become more interconnected, the role of narratives in shaping collective action becomes ever more significant. Behavioral governance, when rooted in transparent practices and shared values, has the potential to guide communities toward sustainable and equitable futures. The challenge, therefore, is not to abandon behavioral governance but to refine its methods to ensure that it supports collaboration, equity, and the pursuit of the common good.

References

- Adler, E., & Bernstein, S. (2005). Knowledge in power: the epistemic construction of global governance. In M. Barnett, & R. Duvall (Eds.), *Power in Global Governance* (pp. 294–318). Cambridge University Press. <https://doi.org/10.1017/CBO9780511491207.013>
- Arendt, H. (1998). *The Human Condition*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226924571.001.0001>
- Ariely, D. (2008). *Predictably Irrational*. Harper Collins.
- Bachrach, P., & Baratz, M. S. (1970). *Power and Poverty: Theory and Practice*. Oxford University Press.
- Battaglio, P. R., Belardinelli, P., Bellé, N., & Cantarelli, P. (2019). Behavioral Public Administration ad fontes: A Synthesis of Research on Bounded Rationality, Cognitive Biases, and Nudging in Public Organizations. *Public Administration Review*, 79(3), 304–320. <https://doi.org/10.1111/puar.12994>
- Battegazzorre, F. (2017). Some reflections on the power to/power over debate. *Journal of Political Power*, 10(3), 274–285. <https://doi.org/10.1080/2158379X.2017.1382169>
- Castells, M. (2011). A Network Theory of Power. *International Journal of Communication*, 5, 773–787. Online: <https://ijoc.org/index.php/ijoc/article/view/1136>
- Chetty, R. (2015). Behavioral Economics and Public Policy: A Pragmatic Perspective. *American Economic Review*, 105(5), 1–33. <https://doi.org/10.1257/aer.p20151108>
- Dahl, R. A. (1957). The concept of power. *Behavioral Science*, 2(3), 201–215. <https://doi.org/10.1002/bs.3830020303>
- De la Cruz Prego, F. (2021). The Role of the State in the 2030 Agenda: An Economic Policy Approach. *Revista de Economía Mundial*, 63, 97–116. <https://doi.org/10.33776/rem.vi63.7149>

- Dowding, K. (2019). *Rational Choice and Political Power*. Bristol University Press. <https://doi.org/10.51952/9781529206340>
- Dowding, K., & Oprea, A. (2024). Manipulation in politics and public policy. *Economics and Philosophy*, 40(3), 685–710. <https://doi.org/10.1017/S0266267124000063>
- Fisher, W. (1989). *Human Communication as Narration: Toward a Philosophy of Reason, Value, and Action*. University of South Carolina Press. <https://doi.org/10.2307/j.ctv1nwbqtk>
- Foucault, M. (1980). Truth and power. In C. Gordon (Ed.), *Power/Knowledge: Selected Interviews and Other Writings, 1972–1977* (pp. 109–133). Pantheon Books.
- Foucault, M. (1997). *Il faut défendre la société: Cours au Collège de France, 1976*. Gallimard/Seuil.
- Foucault, M. (2009). *Security, Territory, Population. Lectures at the Collège de France, 1977–1978*. Palgrave Macmillan. <https://doi.org/10.1057/9780230245075>
- Foucault, M. (2014). The politics of health in the eighteenth century. *Foucault Studies*, 18, 113–127. <https://doi.org/10.22439/fs.v0i18.4654>
- Galizzi, M. M. (2017). Behavioral aspects of policy formulation: experiments, behavioral insights, nudges. In M. Howlett, I. Mukherjee, & S. Fraser (Eds.), *Handbook of Policy Formulation. Handbooks of Research on Public Policy* (pp. 410–429). Edward Elgar Publishing. <https://doi.org/10.4337/9781784719326.00034>
- Gigerenzer, G. (2014). Breast cancer screening pamphlets mislead women. All women and women’s organizations should tear up the pink ribbons and campaign for honest information. *British Medical Journal*, 348, g2636. <https://doi.org/10.1136/bmj.g2636>
- Gofen, A., Moseley, A., Thomann, E., & Weaver, R. K. (2021). Behavioural governance in the policy process: introduction to the special issue. *Journal of European Public Policy*, 28(5), 633–657. <https://doi.org/10.1080/13501763.2021.1912153>
- Grewal, D. S. (2008). *Network Power. The Social Dynamics of Globalization*. Yale University Press. <https://doi.org/10.12987/9780300145120>
- Grimmelikhuijsen, S., Jilke, S., Olsen, A. L., & Tummers, L. G. (2017). Behavioral Public Administration: Combining Insights from Public Administration and Psychology. *Public Administration Review*, 77(1), 45–56. <https://doi.org/10.1111/puar.12609>
- Grüne-Yanoff, T. (2012). Old wine in new casks: libertarian paternalism still violates liberal principles. *Social Choice and Welfare*, 38, 635–645. <https://doi.org/10.1007/s00355-011-0636-0>
- Haas, P. M. (1992). Epistemic communities and international policy coordination. *International Organization*, 46(1), 1–35. <https://doi.org/10.1017/S0020818300001442>
- Halpern, D., & Sanders, M. (2016). Nudging by Government: Progress, Impact, & Lessons Learned. *Behavioral Science & Policy*, 2(2), 53–65. <https://doi.org/10.1177/237946151600200206>
- Hallsworth, M., Egan, M., Rutter, J., & McCrae, J. (2018). Gobierno conductual: Aplicación de las ciencias del comportamiento para mejorar la toma de decisiones de los gobiernos. Behavioural Insights Team. Online: <https://www.bi.team/publications/gobierno-conductual/>
- Hallsworth, M., List, J. A., Metcalfe, R. D., & Vlaev, I. (2017). The behaviorist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics*, 148, 14–31. <https://doi.org/10.1016/j.jpubeco.2017.02.003>
- Han, B.-C. (2005). *Was ist Macht?*. Reclam.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.
- Lehner, M., Mont, O., & Heiskanen, E. (2015). Nudging – A promising tool for sustainable consumption behaviour?. *Journal of Cleaner Production*, 134, 166–177. <https://doi.org/10.1016/j.jclepro.2015.11.086>

- Lukes, S. (1974). *Power: A Radical View*. Macmillan.
- Madrian, B. C., & Shea, D. F. (2001). The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior. *The Quarterly Journal of Economics*, 116(4), 1149–1187. <https://doi.org/10.1162/003355301753265543>
- Miller, P., Kable, J. W., & Lempert, K. M. (2024). Willingness to wait outperforms delay discounting in predicting drinking severity. *Journal of Experimental Analysis of Behavior*, 122(3), 247–258. <https://doi.org/10.1002/jeab.4210>
- Moseley, A., & Thomann, E. (2021). A behavioural model of heuristics and biases in frontline policy implementation. *Policy and Politics*, 49(1), pp. 49–67. <https://doi.org/10.1332/030557320X15967973532891>
- Osman, M., & Thornton, K. (2019). Traffic light labelling of meals to promote sustainable consumption and healthy eating. *Appetite*, 138, 60–71. <https://doi.org/10.1016/j.appet.2019.03.015>
- Parsons, T. (1967). *Sociological Theory and Modern Society*. Free Press.
- Pettit, P. (1996). Freedom as Antipower. *Ethics*, 106(3), 576–604. <https://doi.org/10.4135/9781446215272.n18>
- Porter, T. (2009). Global Governance as Configurations of State/Non-State Activity. In J. Whitman (Ed.), *Palgrave Advances in Global governance* (pp. 87–104). Palgrave Macmillan. https://doi.org/10.1057/9780230245310_5
- Rawls, J. (1993). *Political Liberalism*. Columbia University Press.
- Schön, D., & Rein, M. (1994). *Frame Reflection: Toward the Resolution of Intractable Policy Controversies*. Basic Books.
- Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 4(2), 155–169. <https://doi.org/10.1007/BF01405730>
- Sassen, S. (2006). *Territory, Authority, Rights: From Medieval to Global Assemblages*. Princeton University Press. Online: <http://www.jstor.org/stable/j.ctt7sx98>
- Simon, H. (1978). Rationality as Process and as Product of Thought. *The American Economic Review*, 68(2), 1–16. Online: <https://www.jstor.org/stable/1816653>
- Sunstein, C. R. (2014). Nudging: A Very Short Guide. *Consumer Policy*, 37, 583–588. <https://doi.org/10.1007/s10603-014-9273-1>
- Sunstein, C. R. (2019). *Conformity. The Power of Social Influences*. New York University Press. <https://doi.org/10.18574/nyu/9781479896585.001.0001>
- Thaler, R. H., & Sunstein, C. R. (2003). Libertarian Paternalism. *American Economic Review*, 93(2), 175–179. <https://doi.org/10.1257/000282803321947001>
- Thaler, R. H., & Sunstein, C. R. (2008). *Improving Decisions About Health, Wealth, and Happiness*. Yale University Press.
- Thomann, E., van Engen, N. A. M., & Tummers, L. (2018). The necessity of discretion: A behavioral evaluation of bottom-up implementation theory. *Journal of Public Administration Research and Theory*, 28(4), 583–601. <https://doi.org/10.1093/jopart/muy024>
- Thomas, M. D. (2019). Reapplying behavioral symmetry: public choice and choice architecture. *Public Choice*, 180, 11–25. <https://doi.org/10.1007/s11127-018-0537-1>
- Wartenberg, T. E. (1990). *The Forms of Power. From Domination to Transformation*. Temple University Press.
- White, M. D. (2013). *The Manipulation of Choice. Ethics and Libertarian Paternalism*. Palgrave Macmillan.
- Wilkinson, T. M. (2013). Nudging and Manipulation. *Political Studies*, 61, 341–55. <https://doi.org/10.1111/j.1467-9248.2012.00974.x>