

# **Nudging in times of an environmental crisis**

## **An assessment of the moral permissibility of the City of Cape Town's Water Map**

**Barbara Knoote**

**MA Thesis Applied Ethics**

**Word Count: 17290**

**18/06/2019**

### **Abstract**

The Water Map is a social norms nudging intervention that was implemented by the municipal government of Cape Town with the aim to reduce residents' water consumption during a critical drought in the city. The Water Map made public data about the water use of Cape Town's residents, which raised concerns amongst the residents about privacy and naming and shaming. In this paper, I will assess whether naming and shaming practices were used and whether there was a violation of privacy. Next, I will determine what this means for the moral permissibility of the Water Map. The conclusion from the assessment is that the Water Map was a morally impermissible policy intervention because it violated privacy. This violation of privacy was not justified because it was not based on evidence that showed that a more privacy-respecting social norms nudge would lead to a decrease of the intervention's effectiveness. In the last part of the paper, some lessons for the future of social norms nudging are discussed.

# Table of Contents

<b>Introduction</b>	<b>p. 4</b>
<b>1. Social norms nudging in public policy</b>	<b>p. 12</b>
1.1 What is nudging?	p. 12
1.2 How do governments nudge?	p. 14
1.3 Social norms nudges	p. 16
<b>2. The value of privacy</b>	<b>p. 18</b>
2.1 The value of informational privacy	p. 19
2.2 Information as a tradable good	p. 20
<b>3. Naming and shaming</b>	<b>p. 22</b>
3.1 What is naming and shaming?	p. 22
3.2 Why can naming and shaming be morally problematic?	p. 23
<b>4. An assessment of the moral permissibility of the Water Map</b>	<b>p. 25</b>
4.1 A description of the characteristics of the Water Map	p. 25
4.2 The Water Map: nudge or sludge?	p. 27
4.3 The water Map and the violation of privacy	p. 28
4.4 The Water Map and public shaming	p. 32
4.5 Conclusion	p. 34
<b>5. Lessons for the future</b>	<b>p. 35</b>
<b>Conclusion</b>	<b>p. 38</b>
<b>Bibliography</b>	<b>40</b>

**This page is intentionally left blank**

## Introduction

In 2017 and 2018, The city of Cape Town experienced its worst drought in decades<sup>1</sup>. The dams that the city relies on for its water provision were close to empty, and there was no prospect of rainfall. If the residents of Cape Town would not drastically reduce their water consumption, the city would become the first major city to run out of water completely: this prospect was called ‘Day Zero’. Day Zero referred to the day when city officials would be forced to cut off the normal water supply to most city’s homes. From that day on, residents would be forced to queue for water at one of the 200 water collection points. There, they would receive an allocation of 25 litres of water per person. The collection sites were said to be guarded by police or military personnel. Initially, Day Zero was expected to ‘happen’ on the 16th of April of 2018. Later, it was said to take place on May 11th 2018<sup>2</sup>. Day Zero soon became the talk of the town.

Something needed to change, and action was required from all three levels of government: national government, the provincial government and the municipality of Cape Town. During the drought, these institutions implemented numerous interventions to try and safeguard the city’s water reserves. Luckily, the city ended up succeeding in limiting their water consumption to the bare minimum. That, together with a bit of rainfall, resulted in what is at least a temporary relief of the water crisis. Day Zero was called off. One of the first things that the City of Cape Town did was increasing the water tariffs and putting into place water restrictions<sup>3</sup>. However, when the crisis became increasingly severe, and these interventions did not seem to make a significant impact, the City of Cape Town started experimenting with more innovative interventions. One of these interventions was called the ‘Cape Town Water Map’.

The Cape Town Water Map was an online map that showed the water consumption of households online and was introduced in January of 2018 when the city’s drought was reaching its worst point. Everyone could now go online and see how much water other people in the Cape Town were consuming. Furthermore, the map provided information on collection points for recycled wastewater, and water distribution zones.

The Water Map is a policy intervention that falls into the category of ‘nudging’. In the book *Nudge* by Thaler & Sunstein, nudges are defined as: ‘interventions that steer people in particular directions but that also allow them to go their own way<sup>4</sup>’. More elaborate, they are characterized in the book as such:

---

<sup>1</sup> Pedro M Sousa et al., ‘The ‘Day Zero’ Cape Town drought and the poleward migration of moisture corridors,’ *Environmental Research Letters* 13, no. 12 (December 2018)

<sup>2</sup> ‘Day Zero pushed back from April 16 to May 11,’ *Mail & Guardian*, February 5, 2018, <https://mg.co.za/article/2018-02-05-day-zero-pushed-back-to-may-11>

<sup>3</sup> ‘Level 2 water restrictions imposed on Cape Town residents’, *Enca*, December 2015, <https://www.enca.com/south-africa/level-2-water-restrictions-imposed-cape-town-residents>

<sup>4</sup> Richard H. Thaler and Cass R. Sunstein, 2008, ‘*Nudge: Improving Decisions About Health, Wealth, and Happiness*’, Revised & Expanded edition. (New York: Penguin Books).

*“A nudge, as we will use the term, is any aspect of the choice architecture that alters people’s behaviour in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid. Nudges are not mandates. Putting fruit at eye level counts as a nudge. Banning junk food does not.”<sup>5</sup>*

The incentive of the Water Map was thus to try and reduce people’s water consumption. According to The City of Cape Town, the Water Map would encourage people to save water<sup>6</sup>. The mechanism behind this intervention was reasoned to be twofold. The first mechanism would be that people were being rewarded if they saved water because their house would show up on the map as a ‘green’ house. This ‘rewarding’ mechanism would then lead to normalization and incentivisation of water conservation behaviour<sup>7</sup>. The second mechanism would be to create and make transparent a social norm. These two mechanisms are both widespread phenomena in behavioural economics and nudging. The first one is the effect of reward and social status. The second one is the power of social norms. The Water Map is therefore a nudge that steers people’s behaviour towards low water consumption behaviour.

However, not all residents of Cape Town experienced the Water Map as a positive encouragement to save water, but more so as a platform to name and shame the people who did not do this<sup>8</sup>. By labelling the high water consumers as such, and the low water consumers as such, the map was dividing the residents into groups of ‘good citizens’ and ‘bad citizens’, or ‘environmentalists’ and ‘non-environmentalists’. Residents felt like the city was stigmatizing the group of people that were not doing a good job in saving water. Residents also felt like the city was invading their privacy, as information about their personal life now became public. They felt like the city was crossing lines in publishing this data about them online<sup>9</sup>. As a result of this data being available online, some residents were sincerely frightened that the people who showed on the map as ‘high water consumers’ would be threatened or harassed<sup>10</sup>. These fears and critiques give rise to the question of whether the Cape Town Water Map was a morally permissible policy intervention. In this paper, I will do an ethical assessment of the intervention that I will refer to as the Water Map. In order to do this, I will assess whether a violation of privacy causes the intervention to be morally impermissible. I will also assess whether the Water Map’s use of naming and shaming practices is causing the intervention to be morally impermissible.

<sup>5</sup> Thaler and Sunstein, *Nudge: Improving Decisions About Health, Wealth, and Happiness*, 6.

<sup>6</sup> Sinclair-Smith, K. et al., “City of Cape Town’s Water Map”. *American Water Works Association*, 110, (2018): 65.

<sup>7</sup> Sinclair-Smith, K. et al., “City of Cape Town’s Water Map”, 64

<sup>8</sup> Olivier DW., “Cape Town’s map of water usage has residents seeing red”. *Conversat*. January 2018; <http://theconversation.com/cape-towns-map-of-waterusage-has-residents-seeing-red-90188>

<sup>9</sup> Bobby Jordan, Useful water-saving tool? Or has Cap Town gone dotty?, TimesLive, 15 January 2018. <https://www.timeslive.co.za/news/south-africa/2018-01-15-useful-water-saving-tool-or-has-cape-town-gone-dotty/>

<sup>10</sup> Bobby Jordan, Useful water-saving tool? Or has Cap Town gone dotty?

### **Shifting the focus from autonomy and manipulation**

By focussing on privacy, naming and shaming and the moral permissibility of social norms nudging in my ethical assessment, I am taking a new approach to the ethical assessment of nudges. There are two things in particular that I am stepping away from. The first thing that I will be stepping away from in this ethical assessment, is merely putting my focus on autonomy and manipulation. Most critiques of nudging focus on autonomy and manipulation. There are roughly two sides to this debate. One side believes that nudging is a revolutionary tool that can improve public policy and human wellbeing by helping people to make better decisions. The other side is generally worried about the moral permissibility of nudging because it would be a manipulative way of policymaking that would undermine autonomy. In this paper, I will move away from this debate about autonomy and manipulation.

The first reason not to focus on autonomy and manipulation is that not *all* nudging undermines respect for autonomy and/or is manipulative. The next paragraph will provide the main reasons for this statement. As for the critique that nudging is not respecting personal autonomy enough: because the goal of nudging is to push someone in a certain direction, it is said to undermine a person's personal opinion of what the 'right direction' is. It is, therefore, being criticized as being too paternalistic by many authors. They claim that nudges do not respect individuals as autonomous agents that are perfectly capable of making well-considered decisions. To these criticisms, Sunstein responds by stating that good nudges do not disregard an agent's autonomy. A lot of nudges even promote personal autonomy. He explains: "Consider, for example, an effort to prompt people to make their own choices about what kind of retirement plan they want, by asking them precisely that question when they begin employment<sup>11</sup>. As for the critique of nudging being manipulative: Sunstein responds to this criticism by saying that people cannot escape being nudged by their surroundings. In order to understand this, one has to understand the concept of choice architecture. Choice architecture is the way in which choices are presented to individuals, which impacts decision-making. In a choice architecture, there are many variables, like the number of choices presented, and whether there is a 'default' or not. All these things influence decision-making, and there is no way to escape this. Johnson (2012), also described this in his essay when he explained that neutral choice architecture is an illusion since any way in which a choice is presented has some kind of influence on decision-making<sup>12</sup>. Meaning: nudges only drive the choice architecture towards certain outcomes. In this respect, being defaulted into not being enrolled in a pension program is equally manipulative as being defaulted into enrolment in a pension programme. Therefore, because not all nudges are automatically a threat to autonomy or rely on manipulation, it is not absolutely necessary for me to incorporate these values into my ethical assessment.

---

<sup>11</sup>Cass Sunstein, "The Ethics of Nudging," *Yale Journal on Regulation* 32, no. 2 (2015): 415.

<sup>12</sup> Johnson et al., "Beyond nudges: Tools of a choice architecture" *Marketing Letter* 23, (June 2012): 488, <https://doi.org/10.1007/s11002-012-9186-1>

The second thing that I will be stepping away from in this paper, is seeing all nudges as one and the same technique of policy-making. Usually, when people criticise nudging, they tend to focus on all types of nudges, or on the entire concept of nudging. However, there are many different kinds of nudges that rely on vastly different behavioural insights and methods. When you focus on the moral permissibility of all nudges, you cannot do justice to the unique problems in moral justification that the different kinds of nudges bring about.

That being said, if significant issues arise that demonstrate that the Water Map poses a threat to personal autonomy or relies heavily on manipulation I will of course take that into account in my ethical assessment of the Water Map. Moreover, if the need arises to zoom-out from the specific kind of nudge that I will discuss in this paper to discuss nudging more broadly, I will do so. However, it will not be my main focus.

### **Privacy**

There are three reasons why I will focus on the value of privacy in the ethical assessment of the Water Map. The first and most straightforward reason is that the public debate in South Africa about these interventions demonstrated that the residents of Cape Town experienced a violation of their privacy<sup>13</sup>. A second reason for looking at privacy is that this paper contributes to the ongoing debate between privacy and other values, such as safety or economic gain. This is a debate that has been continuously growing ever since our world has become more and more digital. In a world with more information about individuals than ever before, it is important to re-evaluate this value of privacy, to see if and why it is still a value of importance. Part of this debate also concerns the responsibilities and duties of companies and governments. The third reason for focussing on privacy is that social norms nudges often rely on information sharing, which is a method that is extremely prone to privacy violations. With the increasing use of nudges amongst companies and governments, the use of social norms nudging also increases. It is therefore of great relevance to find out what role the value of privacy plays in the justification of these kinds of nudges.

### **Naming and shaming**

Another concept that I will investigate more closely is the concept of naming and shaming. There is one main reason why I chose to focus on this phenomenon in my ethical assessment of the Water Map.

The reason for focussing on naming and shaming is because it is questionable whether interventions that are based on naming and shaming residents are morally justified. Although it can definitely be a powerful tool to change undesired behaviour, it can also lead to a lot of damage done to the people who are being shamed. Publicly shaming is a strategy that was used as a form of punishment in the past centuries. An example of this is the regular use of pillories in medieval times. Now, however,

---

<sup>13</sup> Olivier DW., "Cape Town's map of water usage has residents seeing red". *Conversat*. January 2018; <http://theconversation.com/cape-towns-map-of-waterusage-has-residents-seeing-red-90188>

it is perceived as an outdated unethical form of punishment, and as a danger to human dignity. Nonetheless, it seems like naming and shaming is making its comeback with the rise of social media. But how can we compare naming and shaming practices from the past that the majority of people frown upon so much, with new naming and shaming practices that take place online?

In my ethical assessment of the moral justification of the Water Map, I will try to answer the question what it means for the moral permissibility of the Water Map if this intervention relies on a naming and shaming method.

It has to be said that there are many other values and principles that play a role in the moral justification of a nudge or any other intervention based on behavioural insights. Unfortunately, I will not be able to scrutinize all of them in this paper. However, there are some values or principles that are of such importance in the field of behavioural insight interventions, that I cannot disregard them completely. In the next paragraphs, I will introduce the concerning values and explain how I will use them in my current reflection and why I have chosen to do so.

### **Effectiveness**

Effectiveness is a value that cannot be overlooked when it comes to policy interventions. Policy interventions are designed and rolled-out with the goal of effectively managing, regulating or changing something regarding human behaviour within a certain society. When an intervention does not effectively do what it is supposed to do, it might not be justified to allocate money, time and other resources to this intervention. However, this does not necessarily mean that the intervention is *morally* unjustified. Though, there are (at least) two scenarios where one can argue that ineffectiveness of an intervention does lead to the moral impermissibility of this intervention.

1. When the intervention made one or more moral compromises that were supposedly being outweighed by the positive outcomes of the intervention.
2. It can be argued that an ineffective intervention is morally problematic when the resources that were put in the interventions could have been allocated to another intervention that would have brought about more wellbeing, happiness, or positive change.

It is for multiple reasons that I am not focussing too much on this value in this paper. The first reason is that these interventions took place in times of a crisis. Therefore, there was little time to research or pilot the intervention. Because the intervention relied on knowledge from previous studies about social norms nudging, we can suppose that there was at least some reason to assume that the intervention would be effective. The second reason not to focus on effectiveness is that effectiveness is also very hard to measure in a time like a crisis when many interventions are all trying to achieve the same goal.

Although I won't be focussing completely on effectiveness, it will still be a value to keep in mind in the ethical assessment.

### **Well-being**

The other value that is of great importance for the moral justification of policy interventions is the wellbeing of the people that it affects. Utilitarians would even claim that well-being is all that matters and that interventions such as this one could hypothetically be weighed by a utilitarian calculus. That is not what I will be arguing in this paper, as I think well-being is not the only value that has to be taken into consideration. I will, however, not claim that it is not an important value, as it is clearly one of the end goals of making good public policy. Nonetheless, it is important to emphasize that policy interventions cannot only be assessed by its outcomes in terms of happiness or wellbeing, but also by the quality of governance and policy interventions.

### **The bigger picture**

The assessment of the moral justification of the Water Map will do more than to just tell us whether the Water Map was a morally permissible policy intervention or not. I argue that the moral scrutinization of this intervention is of greater importance than to just tell us whether this intervention was morally justified or not. The lessons learnt from this policy intervention in Cape Town can serve almost as a tool to help understand certain moral boundaries in the field of nudging, especially with the use of social norms, in public policy. There are multiple reasons why this is important and relevant.

The first reason is that environmental crises are a specific domain of policymaking. Due to time constraint, there is usually a limited opportunity for policymakers to research or debate certain policy interventions. Therefore, it is important to look at these interventions in hindsight to examine if this was really the best way of dealing with the crisis. This knowledge can then be used for future crises.

The second reason for the relevance and importance of this paper is that nudging is quite a new, yet rapidly growing domain of public policy. The first literature on nudging has only been written a little over ten years ago. Ever since, governments have been picking up on the method rapidly on a global scale. However, because there is no long history of nudging, the governments that use nudges do not have a lot of knowledge from the past to fall back on. Nonetheless, we see governments and companies using it more and more often. Furthermore, it is especially interesting to examine social norms nudging because research suggests it can be a powerful tool to change decision-making. It is exactly this kind of a tool that we need if we want to combat environmental crises. Environmental crises are often (at least partly) collective action problems, meaning that individuals on large scale have to change their behaviour in order to create a better outcome for everyone. If social norms nudging really incentivises individuals to change their behaviour, it could be a very useful tool in many environmental problems that are driven by human behaviour. However, if this is the case, it is important to see if it is morally permissible to continue this method of policymaking.

The third reason why this ethical assessment is relevant for the future is that it focuses on a policy intervention where a governmental organisation made use of data sharing. This will continue to

be relevant for future research because data sharing is currently widely debated, both in commercial contexts as in government contexts. Because there is more data available about individuals than ever, and it seems that this data is worth something to organisations, it is important to review what certain organizations are allowed to do with that data and what they are not allowed to do with it.

The three arguments above demonstrate that the relevance of the Cape Town water crisis and the Water Map intervention is multileveled. On the one side, the crisis depicts a very unique and distinctive environmental problem with a unique policy intervention to combat it. On the other side, the water crisis is almost an exemplifying case of one of the first instances of an urgent environmental crisis that the planet might experience more of in the future. Environmental crises and extreme weather conditions, such as drought, are predicted to increase in the future due to climate change. This entails that the planet and its residents will be facing more environmental crises in the future. How we went about combating environmental crises in the past, and whether the interventions that have been rolled out to combat them are morally permissible is relevant information to take into account for future scenarios.

In this paper, I will try to answer the question of whether the social norms nudge that the city used for the water crisis was morally permissible. The overall structure of the paper takes the form of this introduction, five chapters and a conclusion. In Chapter 1, I will introduce the technique ‘nudging’. I will explain what it is and why governments use it as a policy strategy. I will then introduce what social norms nudges are and how they work. The second Chapter is concerned with the value of privacy. In this chapter, I will discuss the importance of privacy, the different accounts of privacy, and the concept of information as a tradable good. The third Chapter will focus on the technique of naming and shaming. I will both discuss the incentives for naming and shaming, as well as its moral permissibility. In the fourth Chapter, I will turn to the ethical assessment of the Water Map, where I will answer the main research question: ‘Was the Water Map a morally permissible policy intervention?’ In order to answer this question, this paper will focus on three criteria. The first thing that I will examine, is whether the Water Map classifies as a ‘nudge for good’ rather than a ‘sludge’, as defined by author Richard Thaler. The second part of the ethical assessment will focus on the violation of privacy. The third part will look into naming and shaming practices. Based on the conclusions that will be drawn from those three part, the main question will be answered.

Concretely, the questions that I will be answering in Chapter four are the following:

1. Was the Water Map ‘nudging for good’?
2. Did the Water Map violate the resident’s privacy?
  - If so, does this invasion of privacy lead to a moral impermissibility of the Water Map?
3. Did the Water Map use the technique of naming and shaming?
  - If so, does the use of this technique make the Water Map morally impermissible?
4. Was the Water Map a morally permissible policy intervention?

The fifth and last Chapter will focus on lessons for the future of social norms nudging in public policy, and social norms nudging in environmental crises. Lastly, the conclusion will provide an overview of the main findings of this paper, as well as some implications.

## Chapter 1: Socials norms nudging in public policy

In the introduction, I have given a short overview of the drought that struck Cape Town and the behavioural policy intervention that the city implemented during the most crucial period of the drought. In this Chapter, I will provide a bit more context on nudging and social norms nudging in public policy. The first section will introduce the concept of nudging and will introduce the three principles of ‘nudging for good’. The second section will go more into details about how governments use nudges to steer behaviour a certain way. In the third section, the social norms nudge is introduced, which is the specific kind of nudge that will be discussed in this paper.

### 1.1 What is nudging?

Whether you are aware of it or not, you are being influenced by nudges on a daily basis. We buy more of the things that lay on eye-level in the supermarkets<sup>14</sup>. We are more likely to end up as an organ donor if we are required to opt-out in case we do not want this<sup>15</sup>. We also choose the stairs over the lift more often if the stairs have been colourfully painted<sup>16</sup>.

Nudges are interventions that steer people in particular directions but that also allow them to go their own way<sup>17</sup>. ‘Normal’ informative policies often assume that people are completely rational beings, who perform economic rational behaviour<sup>18</sup>. In reality, however, human decision-making is not rational, but often subject to biases<sup>19</sup>. Nudges take into account the way how our decision-making is influenced by internal and external factors. Nudge theory, therefore, relies on insights from behavioural economics, ‘behavioural insights’ (BI). Figure 1 provides some examples of Behavioural Insights.

---

<sup>14</sup> Amy L. Wilson, Elizabeth Buckley, Jonathan D. Buckley and Svetlana Bogomolova. “Nudging healthier food and beverage choices through salience and priming. Evidence from a systematic review,” *Food Quality and Preference* 51, (July 2016): 49

<sup>15</sup> Eric Johnson and Daniel Goldstein, “Default and Donation Decisions,” *Transplantation* 78, (December 2004): 1716

<sup>16</sup> Matthew McLaughlin, John Bellettiere and Natasha Bliss, “One step at a time: Simple nudges can increase lifestyle physical activity”, (October 2017): DOI: 10.13140/RG.2.2.32064.23042

<sup>17</sup> Thaler and Sunstein, *Nudge: Improving Decisions About Health, Wealth, and Happiness*, 6.

<sup>18</sup> Ursula Hansen and Ulf Schrader, “A Modern Model of Consumption for a Sustainable Society,” *Journal of Consumer Policy* 20, (January 1997): 455

<sup>19</sup> Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY, US: Farrar, Straus and Giroux.

Loss aversion

One of the most well-known insights from behavioural sciences is called ‘loss aversion’. Loss aversion is a phenomenon that describes a human’s tendency to prefer not losing something over gaining that same thing. Studies have found that people who lose a certain amount of money will experience a bigger loss of satisfaction than the amount of satisfaction that people who gain the same amount of money.

The default mechanism

The default mechanism is a mechanism that explains that human beings often stay with their default option, which is the option that will apply when someone does not actively make a choice.

Status quo bias

Status quo bias refers to the phenomenon that people prefer things to stay the same and to sticking with a decision made previously

IKEA effect

The tendency for people to place a disproportionately high value on objects that they partially assembled themselves, such as furniture from IKEA, regardless of the quality of the end result.

Figure 1<sup>20</sup>.

Nudges take into consideration these behavioural insights and use them to steer decision-making in a particular direction. There are roughly four categories of nudges<sup>21</sup>. The first category relies on the simplification and framing of information. The second category works with changes in the physical environment. The third relies on making changes to the default option. The fourth and last one is the category that uses social norms comparisons. The kind of nudges that I will discuss in this paper are the social norms nudges.

In order to ensure that nudges are being used in an ethical way that helps people making better decisions, the ‘fathers of nudging’ Richard Thaler and Cass Sunstein propose three principles for ‘nudging for good’. They claim that if the principles are not being respected, we have not to speak of a nudge, but a sludge. A sludge is a nudge-like behavioural intervention that does not succeed in encouraging behaviour that is in line with the individual’s best interest<sup>22</sup>. This means that it is either not encouraging that behaviour at all, or that it is encouraging a certain behaviour but that behaviour is not the behaviour that is in line with the individual’s best interest.

<sup>20</sup> Gigerenzer, G. (2018), The bias bias in behavioral economics. *Review of Behavioral Economics*, 5(3-4), 303-336.

<sup>21</sup> Matthias Lehner, Oksana Mont and Eva Heiskanen, “Nudging - A promising tool for sustainable consumption behaviour?”, *Journal of Cleaner Production*, 135 (2016): 166-177

<sup>22</sup> William Hageman, “When are nudges acceptable? Influences of beneficiaries, techniques, alternatives and choice architects,” *Linköping Studies in Behavioural Science*, 213, (2018)

The three principles for ‘nudging for good’:

1. The nudge should be transparent and never misleading.
2. It should be easy to opt out of the nudge (for example, in a single click).
3. The behaviour being encouraged should improve the welfare of those being nudged.

Thaler and Sunstein argue that interventions that are nudge-like but do not respect the three principles are in fact sludges. Sludges are problematic in their moral justification because they are not ‘nudging for good’, but for other reasons (mostly financial reasons). Thaler and Sunstein argue that ‘nudging for good’ is in fact the only legitimate reason to nudge<sup>23</sup>.

## 1.2 How do governments nudge?

In the past, nudges were mostly seen as good marketing strategies. Now, governments are incorporating nudge theory into many fields of public policy. The reason for the growing popularity of nudging in public policy is that it has proven a very efficient and cost-effective way of intervening with a certain behaviour. Behavioural sciences and insights from psychology and behavioural economics are demonstrating with their research that nudges are often way more effective for changing behaviour than more old-fashioned policy measures such as restrictions or sanctions. The growing amount of literature allows governments to customize a behavioural insight to a nudging intervention of their preference. This evidence-based way of governing a city or country is becoming more and more popular, and the number of people working in behavioural insights is growing rapidly.

One example that demonstrates how countries use a behavioural insight in public policy is the decremental penalty points system for driving offences that Croatia, Bulgaria, France, among others, have implemented<sup>24</sup>. These countries are now using the insight that people are generally averse to losing something (loss aversion) to tackle reckless driving by its citizens. When the driver loses all points, their driving license will be revoked. According to the concept of loss aversion, people are more motivated by the incentive of not wanting to lose something they already have (a point), than gaining something that they do not have already (for example with a point-saving method). In Switzerland, the government is using the insight of defaults to switch from nuclear power sourced energy to more renewable energy. Research has found that although the majority of the people support green electricity (and would also be willing to contribute financially to a greener development of the electricity mix),

---

<sup>23</sup> Richard H. Thaler, “Nudge not Sludge”, *Science* 361 (august 2018): 431-432,

<sup>24</sup> Joana Sousa Lourenço, Emanuele Ciriolo, Sara Rafael Almeida, and Xavier Troussard; “*Behavioural insights applied to policy: European Report 2016*”: page 31, EUR 27726 EN; doi:10.2760/903938

they often stay with the default electricity product offered by their provider<sup>25</sup>. In several cities, a greener electricity mix was then made the default option for citizens, although they could still easily change to a cheaper tariff containing nuclear power sourced energy. Most people decided to stay with the new default option, resulting in large increases of green energy use. In the city of Rorschach, the intervention causes an increase in green energy from 3.6% to 93.2%<sup>26</sup>.

Another well-known example of the power of default options is organ donor registration. Johnson and Goldstein demonstrated that an opt-in system for organ donation leads to a society where much fewer people are registered as an organ donor than an opt-out system, See Figure 2. It also leads to a society where only a small percentage of the people who are willing to be a donor actually register as a donor<sup>27</sup>. This goes against an assumption that is often being made, that humans make conscious decisions about their life all the time. Of course, both the opt-in system as well as the opt-out system have some serious ethical consequences. The opt-in system often leads to too few organ donors, which causes long waiting lists and unnecessary deaths for people who are in need of an organ. It also leads to a big group who are willing donors but who are not registered. The opt-out system, however, leads to a misclassification where people become organ donors against their wishes. This example of organ donor registration demonstrates that changing the default mechanism can be a powerful tool when there is inaction<sup>28</sup>

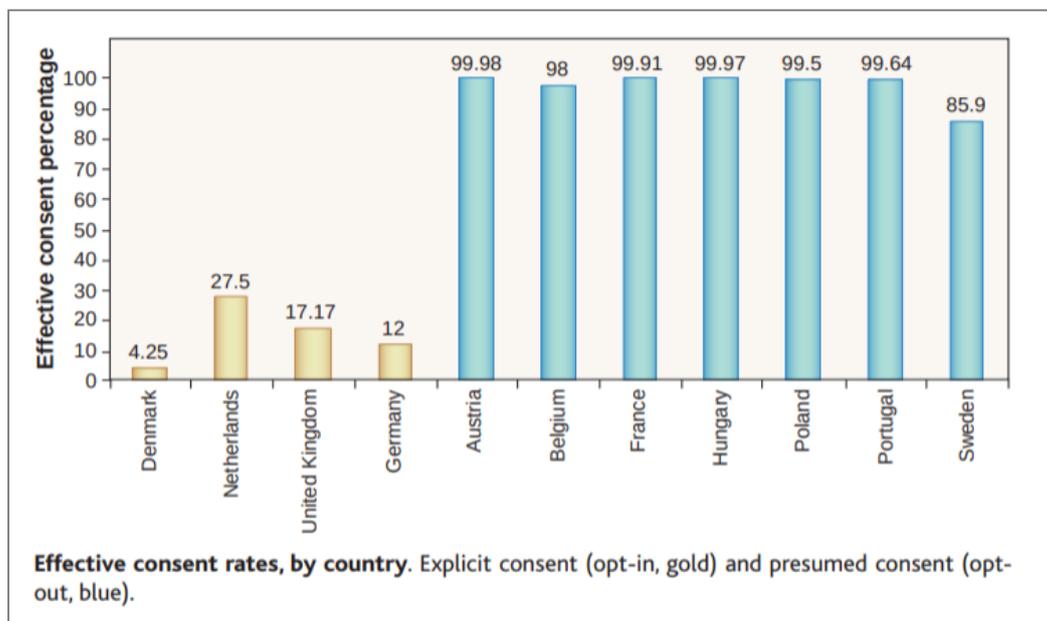


Figure 2

<sup>25</sup> *ibid.*, 22.

<sup>26</sup> *Ibidem.*

<sup>27</sup> Johnson, E. J., & Goldstein, D. G. (2003). "Do defaults save lives?" *Science*, 302, 1338-1339

<sup>28</sup> Samson and Ramani, 2018, "Finding the right nudge for your clients", *Investment News*. Retrieved from <https://www.investmentnews.com/article/20180827/BLOG09/180829939/finding-the-right-nudge-for-your-clients>.

### 1.3 Social norms nudges

Social norms are the informal rules that govern behaviour in groups and societies<sup>29</sup>. Social norms can have a strong effect on behaviour and decision-making. Because of this strong effect, nudges can use these social norms to steer behaviour and decisions in a certain way. In his 2008 publication ‘Nudge, improving decisions about health, wealth and happiness’, Richard Thaler summarizes the effect that social norms have: “Recall that people like to do what most people think it is right to do; recall too that people like to do what most people actually do.”<sup>30</sup>

An example of a strong social norm, and the way it influences behaviour is the silence in libraries<sup>31</sup>. When the power of social norms is being used in behavioural intervention, with the aim to steer behaviour and decision-making a certain way, we can speak of ‘social norms nudging’. Social norms nudges have been proven to be successful in numerous studies. Most of these studies showed a significant change in behaviour when the participants were told what to do in a particular situation using social norms. One study, for example, tested whether hotel guests would be more likely to reuse their towel when they were informed with the percentage of hotel guests that participates in this behaviour already. The sign that the control group saw was a standard message, saying: ‘Help save the environment. You can show your respect for nature and help save the environment by reusing your towels during your stay’. The second group saw this same text, but with an additional descriptive norm message: “Almost 75% of hotel guests participated in the program by using their towels more than once”. The study found that adding a social norms message to the normative message increased the desired behaviour with 9 percentage points (an overall increase of 25,6%).

Research has also found that the nudge was especially powerful when the majority of the people underestimate the proportion of people that engage in the desired behaviour. For example: When a group of taxpayers received a letter that said that 90% of taxpayers paid taxes promptly, this group of people was significantly more likely to pay their taxes in time than the people who had just received a letter reminding them to pay their taxes<sup>32</sup>. This significant effect can be explained by the fact that people had an incorrect perception of the percentage of people that actually paid their taxes in time. Therefore, telling or showing people that the desired behaviour is more common than they think can drive their own behaviour towards the desired behaviour.

The current field of research about social norms nudging is all about effectiveness. Because this kind of policy design is relatively new, there is still a lot of research needed to prove what kind of social

<sup>29</sup> Bicchieri, Cristina, Muldoon, Ryan and Sontuoso, Alessandro, "Social Norms", *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition)

<sup>30</sup> Thaler and Sunstein, *Nudge: Improving Decisions About Health, Wealth, and Happiness*, p.182

<sup>31</sup> Aarts and Dijksterhuis “The silence of the library: Environment, situational norm, and social behaviour”, *Journal of Personality and Social Psychology* 84 (2003) :18–28.

<sup>32</sup> Michael Hallsworth, John A. List, Robert D. Metcalfe and Ivo Vlaev, “The Behaviorist As Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance,” *Journal of Public Economics* 148 (April 2017): 14-31

norms nudges work and that kind does not. Current findings suggest that social norms nudging can be fruitful in environmentally-conscious consumption behaviours<sup>33</sup>. It has also been proven very effective in behaviour and decisions that concern health<sup>34</sup>. Overall, research is showing that social norms nudges can effectively change behaviour, without imposing behavioural change upon people<sup>35</sup>.

Social norms nudging is an upcoming field of research and policy design. Many articles describe particular social norms nudges and analyse their effectiveness. However, in this paper, we will shift the focus from the effectiveness of social norms nudging to the moral permissibility of this practice. Of course, all social norms nudges are different, because there are different ways of how to inform people with this data on social norms. The study that nudged people into reusing their towel provided generic statistics about the percentages of people that chose to reuse their towel. The information was presented in a very private way because you can't actually know who these people are and which of them did reuse their towels and which of them did not. The Water Map took a different approach because it did not provide statistical information, but a map where certain types of behaviour could be linked to the specific address of the people who were actually 'doing that behaviour'. When social norms nudges are more personalized and more public they are more of risk of violating privacy and naming and shaming. Therefore, it is important to keep in mind that different types of social norms nudges make use of different ways of information sharing. The conclusions that will be drawn in the ethical assessment of the Water Map are therefore not directly applicable to other social norms nudges because different types of social norms interventions actually rely on different strategies and designs.

---

<sup>33</sup> Matthias Lehner, Oksana Mont and Eva Heiskanen, "Nudging - A promising tool for sustainable consumption behaviour?", *Journal of Cleaner Production* 134 part A (October 2016): 176

<sup>34</sup> Theresa M Marteau, David Ogilvie, Martin Roland, Marc Suhrcke and Michael P Kelly, "Judging nudging: can nudging improve population health?", *BMJ* 342 (January 2011):228

<sup>35</sup> Berkowitz, "Applications of social norms theory to other health and social justice issues", *In The social norms approach to preventing school and college age substance abuse: A handbook for educators, counselors, and clinicians.* (2003), San Francisco.

## Chapter 2: The value of privacy

In the previous chapter, I introduced the concept of social norms nudging and the broader notion of using nudges in public policy. I explained how the Water Map's unique design gives rise to a debate about the moral permissibility of this nudge and the value of privacy. In this chapter, I will introduce the value of privacy. I will dig into the literature about privacy, after which I will set forth some values that are often conflicting with privacy. This chapter will serve as a basis for chapter 4. In that Chapter, I will use this basis to do an ethical assessment of the Water Map.

First of all, it is important to determine what I will speak about when I speak about privacy since there are multiple conceptualizations of privacy. An important distinction to make is the difference between constitutional privacy and informational privacy. Constitutional privacy, also referred to as decisional privacy, refers to the freedom to make one's own autonomous decisions without interference<sup>36</sup>. Although this kind of privacy is of fundamental importance, I argue that it is not of major relevance in the current case. The severity of the water crisis and the high water consumption of Cape Town's residents were causing a situation that would potentially develop into a really dangerous scenario for all the residents of the city. Interfering with the water consumption of the residents through public policy was, therefore, a very necessary and responsible thing to do. Residents were also free to do with their 'share' of water what they wanted. That they had to stick to a maximum amount of litres per person per day, was a necessary boundary that has little to do with the freedom to make autonomous decisions without interference.

The second kind of privacy, however, is very relevant to the policy interventions that took place in Cape Town during the drought. Informational privacy refers to control over accessing information about oneself. This is the kind of privacy that a lot of Cape Town's residents were concerned with. Informational privacy is the kind of privacy that I will be referring to in this paper.

When wanting to get a better understanding of the current case, it is important to have a good understanding of the value of privacy, but also of the reasons to compromise on informational privacy. In the next section, I will go deeper into the value of privacy. In the following section, I will set forth some reasons for compromising on privacy.

---

<sup>36</sup> Jeroen van den Hoven, Martijn Blaauw, Wolter Pieters, and Martijn Warnier, "Privacy and Information Technology", *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), Edward N. Zalta (ed.), URL <<https://plato.stanford.edu/archives/sum2018/entries/it-privacy/>>.

## 2.1 The value of informational privacy

Amongst the most influential literature when it comes to informational privacy, is Samuel Warren's and Louis Brandeis' essay: 'The Right To Privacy'<sup>37</sup>. In this essay, the authors establish the concept of privacy that is now known as 'control over information about oneself'. This essay did not immediately lead to an implementation of a right to privacy by law. It did, however, start the debate around what a right to privacy would exactly entail. Fifty years later, when the discourse around privacy had been developed further, William Prosser described four privacy rights that he claimed had arisen from different interests in privacy:

1. Intrusion upon [ a person's] seclusion or solitude, or into his private affairs.
2. Public disclosure of embarrassing private facts about [a person].
3. Publicity which places [a person] in a false light in the public eye.
4. Appropriation, for the defendant's advantage, of [a person's] name or likeness<sup>38</sup>.

The importance of the value of privacy is something that has been discussed profoundly in literature over the past decades. Often, different accounts of the value of privacy are placed in one of two categories: reductionists and non-reductionists. The reductionist account of privacy holds that privacy can be valuable but only because it can lead to other valuable things. These things could, for example, be autonomy, security, democracy, utility or economic value<sup>39</sup>. Privacy then becomes valuable because it causes an increase in other valuable goods.

On the other side, there are the non-reductionists, who argue that privacy is valuable because of its inherent value. They claim that there is something about privacy that makes it inherently a valuable and desirable good. There are multiple underlying reasons for philosophers to come to the conclusion that privacy is inherently valuable. One, widespread, view is that privacy is inherently linked to human dignity. In his essay 'privacy as an aspect of human dignity', Bloustein shares his view that privacy protects the non-infringement of personality. This personality that is being protected is one's being, and it requires dignity, integrity, autonomy and independence<sup>40</sup>. A number of authors defend the view that privacy is necessary to have intimate relationships with each other. The privacy that people have to control information about oneself is crucial in establishing these intimate relationships. They argue that the reliance on intimacy on privacy would be (one of) the reason(s) why privacy is valuable<sup>41</sup>. Gerstein (1978), for example, derives this viewpoint from his belief that people require intimacy without

<sup>37</sup> Warren, Samuel D. & Louis D. Brandeis, "The Right to Privacy", *Harvard Law Review*, 4 (1890): 193–220.

<sup>38</sup> William Prosser, "Privacy", *California Law Review*, 48(1960): 383–423

<sup>39</sup> Thomson, Judith Jarvis, "The Right to Privacy", *Philosophy and Public Affairs*, 4 (1975): 295–314.

<sup>40</sup> E Bloustein "Privacy as an Aspect of Human Dignity: An Answer to Dean Prosser", *New York University Law Review*, 39, (1964): 962–1007

<sup>41</sup> Cohen, *Regulating Intimacy: A New Legal Paradigm*, (2002) Princeton: Princeton University Press

observation for them to have relationships without shame<sup>42</sup>. Other authors, however, defend this same view but do not restrict relationships to intimate relationships. They argue that privacy is necessary to have any kind of social relationship. Rachels (1975) argues that it is our ability to control information about us and access to us that controls our relationship with others<sup>43</sup>. Solove takes this claim a step further by arguing that privacy is not only valuable for the individual, but also for the society as a whole. Solove (2008) states:

*“By understanding privacy as shaped by the norms of society, we can better see why privacy should not be understood solely as an individual right.... Instead, privacy protects the individual because of the benefits it confers on society.” Moreover, “the value of privacy should be understood in terms of its contribution to society<sup>44</sup>”*

By claiming that privacy is not merely an individual right, but a value of society Solove argues that societies need respect for privacy. Without it, a suffocating society is created<sup>45</sup>.

## **2.2 Information as a tradable good.**

In practice, we see that the right to informational privacy is often not an absolute right. There are multiple reasons for governments to compromise on informational privacy of their residents or citizens. Reasons for compromising are often to ensure security, to accomplish more transparency, to promote scientific progress, or to produce more economic gain. Our personal information is thus often ‘traded’ for other values. However, not everyone agrees with the fact that our personal information should be considered a tradable good. Roessler describes how tradability of personal information can become harmful and dangerous for people and their social relations<sup>46</sup>. This is in line with Rachel’s theory that describes that informational privacy is crucial to establish relationships with other individuals.

Another problem that arises when informational privacy is compromised on is the problem that Solove describes in his paper ‘nothing to hide, the False trade-off between privacy and security’. In this paper, Solove argues that a trade-off between the two values often leads to a total misbalance with the result that both values are being compromised<sup>47</sup>. Solove argues that the debate relies on false assumptions about the relationship between privacy and safety, like the assumption that privacy cannot be protected without undue cost to security, and the assumption that privacy and security are mutually exclusive<sup>48</sup>. Another theory describes the relationship between privacy and safety as a ‘pendulum’-

<sup>42</sup> Gerstein, R., 1978, “Intimacy and Privacy”, *Ethics*, 89: 76–81

<sup>43</sup> Rachels, “Why Privacy is Important”, *Philosophy and Public Affairs*, 4 (1975): 323–33

<sup>44</sup> Solove, *Understanding Privacy*, (Cambridge, MA: Harvard University Press, 2008).

<sup>45</sup> Solove, *Understanding Privacy* (2008)

<sup>47</sup> Solove, *Understanding Privacy* (2008)

<sup>48</sup> Solove, *Understanding Privacy* (2008)

mechanism. In a pendulum mechanism, the weight or importance can move between two values, meaning that in times of safety crisis or threat, the pendulum should swing towards the side of security. When the need is no longer there, the pendulum will swing back to rights and privacy. A downside of the pendulum-mechanism is that in practice, the pendulum often does not swing back to the sight of privacy because of increasing standards of security.

## Chapter 3: Naming and shaming

In the previous chapter, I introduced the value of privacy. I set forth some well-known theories of the value of privacy, as well as some common reasons to compromise on privacy. In this chapter, I will introduce the concept of ‘naming and shaming’, and ‘public shaming’. I will dig into the literature about naming and shaming, and I will set forth three reasons for why and how this practice is being used. I will then go into the reasons for why the practice can be morally problematic. Similarly to Chapter 2, this chapter will serve as a basis for chapter 4 and chapter 5. In those chapters, I will use this basis to do an ethical assessment of the Water Map.

### 3.1 What is naming and shaming?

Public shaming does not take place the same way anymore now, compared to for example the middle ages where it was standard procedure to put lawbreakers in the pillory in a public space<sup>49</sup>. Shaming in this time was used truly as a way of imposing punishment. Later, in the early 1800s, most judicial systems in western countries started to move away from using these kinds of methods, and disciplinary power became more customary. Naming and shaming, or public shaming, is the act of publicly outing information about an individual or group of individuals, which seeks to induce shame in that person (or at least express a judgement that the person ought to feel ashamed of themselves)<sup>50</sup>.

There are generally three incentives for naming and shaming: to punish, to inform and to criticise<sup>51</sup>. Advocates of the first incentive for, to punish, argue that the practice can be justified for several reasons. The most straightforward one is that it can serve as a deterrent. The negative consequences that result from being publicly shamed, such as a loss of opportunities, reputation and status, can become an incentive for other individuals not to engage in the same illegal or undesirable act<sup>52</sup>. Critics, however, say that this might not be the case at all and that some individuals will be more likely to engage in the illegal act when it is given more recognition<sup>53</sup>. Another criticism against naming and shaming with the goal of punishing is that the public response cannot be predicted and that individuals might be disproportionately punished for their wrongdoings.

The second incentive is to inform people with information. For example, giving out the names and addresses of sex offenders in a certain neighbourhood. If this information is of great importance to other individuals, it might be morally justified to name and shame the other individual. However, a

---

<sup>49</sup> W. Andrews, *Old Time Punishments* (New York: Dorset Press 1890), page 175

<sup>50</sup> Julian Petley, *Media and Public Shaming: Drawing the Boundaries of Disclosure*, Reuters Institute for the Study of Journalism, University of Oxford.

<sup>51</sup> Jacob Rowbottom, “To Punish, Inform, and Criticise: The Goals of Naming and Shaming”, in *Media and Public Shaming: Drawing the Boundaries of Disclosure*, Reuters Institute for the Study of Journalism, University of Oxford. 1-19

<sup>52</sup> Jacob Rowbottom, “To Punish, Inform, and Criticise: The Goals of Naming and Shaming,” 8

<sup>53</sup> Dan M. Kahan, “*What’s really wrong with shaming sanctions*”, Faculty Scholarship Series. 102 (2006): 24

critique of naming and shaming with the goal of informing is that the media that engages in the naming and shaming often dramatizes the story, resulting in the information not being informative at all, but mostly sensational.

The third incentive to public shaming is to criticise a person's conduct<sup>54</sup>. Criticising and informing often goes hand-in-hand, and is hard to distinguish from one another. When a media outlet informs their readers about certain conduct of an individual or organization, they often also share an opinion about that conduct, which can lead to criticising. If the conduct does not match the values of the authors and/or readers, this information can lead to large-scale disapproval of the conduct, which can induce shame, or a feeling that the person ought to be ashamed of themselves.

Even though Rowbottom's three incentives for public shaming all have unique features and can be used individually from each other, Rowbottom argues that the three incentives are in practice difficult to separate and all may arguably be served by the same media campaign. In Chapter 4, I will reflect on the three incentives for public shaming and try to see what incentives were used by the makers of the Water Map. I will then assess the moral permissibility of these incentives.

### The case of Justine Sacco

On the 20th of December in 2013, Justine Sacco made the mistake to put online a 'funny' tweet that was not so funny after all, saying: *'Going to Africa. Hope I don't get AIDS. Just kidding. I'm white!'*

While Justine was on the plane, her tweet went viral. Millions of people on Twitter demanded that Justine should be fired from her job the second that the plane hit the ground. A hashtag was created: #HasJustineLandedYet. When Justine arrived in Africa, she was awaited by people taking pictures of her and posting it online, while she found out that she had been fired from her job whilst being on a plane, unable to respond to any of it.



### 3.2 Why can naming and shaming be morally problematic? (when does it become morally problematic?)

Most of the public shaming practices nowadays take place online and is often done by the media or by individuals. I argue that there are several reasons why online naming and shaming can be morally problematic. The first reason is that public shaming online can lead to a kangaroo court that takes no notice of standards of law or justice. The second reason is that the consequences from naming and shaming are often disproportionate to the conduct. The third reason is that public shame can lead to severe psychological problems such as anxiety or depression. To illustrate these three problematic consequences, I will use a case from John Ronson's book: *'So You've Been Publicly Shamed'*<sup>55</sup> (See: The case of Justine Sacco).

<sup>54</sup> Jacob Rowbottom, "To Punish, Inform, and Criticise: The Goals of Naming and Shaming," 9

<sup>55</sup> Jon Ronson, *So You've Been Publicly Shamed* (New York, Riverhead Books, A member of Penguin Group (USA), 2015)

A majority of the online public shaming that we see online is the shaming of ‘ordinary’ people for minor crimes or non-criminal conduct that people disapprove of. Often, the person engages in behaviour that is perceived as negative and that is also considered ‘not accordant to the norm’. The reason for naming and shaming these people is often to criticise, although punishment or informing could well be motives too. Though, whatever the reason for public shaming might be, it often occurs that online public shaming leads to a ‘kangaroo court’ that takes no notice of standards of law or justice, which is the first reason for why naming and shaming can become morally problematic. In the case of Justine Sacco, people started harassing her online. They found all the information that you can possibly find about someone online. Her flight number was traced, hence the hashtag #HasJustineLandedYet came into existence. Millions of people all over the world were now following Justine’s flight to track whether she had landed yet. Some people even went to the airport to wait her up only to take pictures of her whilst she connected her phone to the Wi-Fi network, to see that her Tweet went viral and to find out that she was now amongst the most hated persons on the internet. These were all consequences from online shaming for something that is not even considered criminal conduct in most countries. The online harassment, and the invasion of Justine’s privacy, were consequences from online shaming that did not take notice of standards of law or justice.

The second reason for naming and shaming to be morally problematic, is that the consequences from naming and shaming are often disproportionate to the conduct. In the case of Justine Sacco, one of the consequences was that she was instantly fired from her job before she was even given notice. Because public shaming (even for minor non-criminal wrongdoings) can place a person in extreme negative light, one of the consequences can be that the status of that person is diminished to the point that it becomes really difficult to participate in society. Certain things, like finding a job or a place to live, become close to impossible because of the negative status of that person. The extend of scapegoating to this person is often not in proportion to the wrongdoing that they actually did. Because of these disproportionate consequences, public shaming can be morally problematic.

The third reason why naming and shaming can be morally problematic is that the consequences from online public shaming can lead to severe psychological problems with the person that has been shamed. After being severely shamed online, many people show symptoms of depression, insomnia, PTSD, severe anxiety and even suicidal thoughts<sup>56</sup>. These psychological problems are often an effect of the first and second reason that I gave above.

In the next Chapter, I will reflect back on the three incentives for naming and shaming, as well as on the three reasons why it can be morally problematic. I will then determine what is applicable to the Water Map and what is not in my assessment of the moral permissibility of Water Map.

---

<sup>56</sup> Jon Ronson, *So You've Been Publicly Shamed* (New York, Riverhead Books, A member of Penguin Group (USA), 2015)

## Chapter 4: An assessment of the moral permissibility of the Water Map

In the previous Chapters, I have introduced the topic and the two values that I will test my case against. This Chapter will provide argumentation for the assessment of the moral permissibility of the Water Map. The first section (4.1) will give an elaborate description of the characteristics of the Water Map, because this will provide the reader with the necessary information to understand later argumentation. Section 4.2 will provide argumentation for whether the Water Map classifies as a nudge or a sludge, according to the three principles for ‘nudging for good’. If the Water Map classifies as a sludge, this might already be morally problematic for the permissibility of the intervention because the nudge is not actually encouraging behaviour that is in line with the individual’s best interest<sup>57</sup>. In section 4.3, I will assess whether the Water Map did or did not violate the privacy of Cape Town’s residents. If I have to conclude that the intervention indeed violated privacy, I will then go ahead and examine whether this has any implications for its moral permissibility. In section 4.4, I will examine whether the Water Map named and shamed people. I will then see what the implications of that conclusion are to the moral permissibility of the Water Map. Section 4.5 will provide conclusions about the moral permissibility of the Water Map.

### 4.1 A description of the characteristics of the Water Map

In January of 2018, the city of Cape Town introduced the Cape Town water consumption map. The main motive for implementing this intervention was a study from 2014 that showed that households would drastically reduce their water use when being given the information that their usage is above average in their neighbourhood<sup>58</sup>. However, many residents of Cape Town were not too happy about the implementation of the Water Map. Various newspaper articles and interviews indicate that there were many concerns. Nonetheless, the municipal government made clear that their motives were positive, and positive only. When the mayor introduced the online tool, she talked about wanting to ‘paint the town green’<sup>59</sup>. ‘Green’ in how it is used here can refer to the green dots that the Water Map displays, as well as green environmental choices by Cape Town’s residents. Figure 3 shows a picture of the Water Map. Figure 4 shows the caption that was provided with the Water Map.

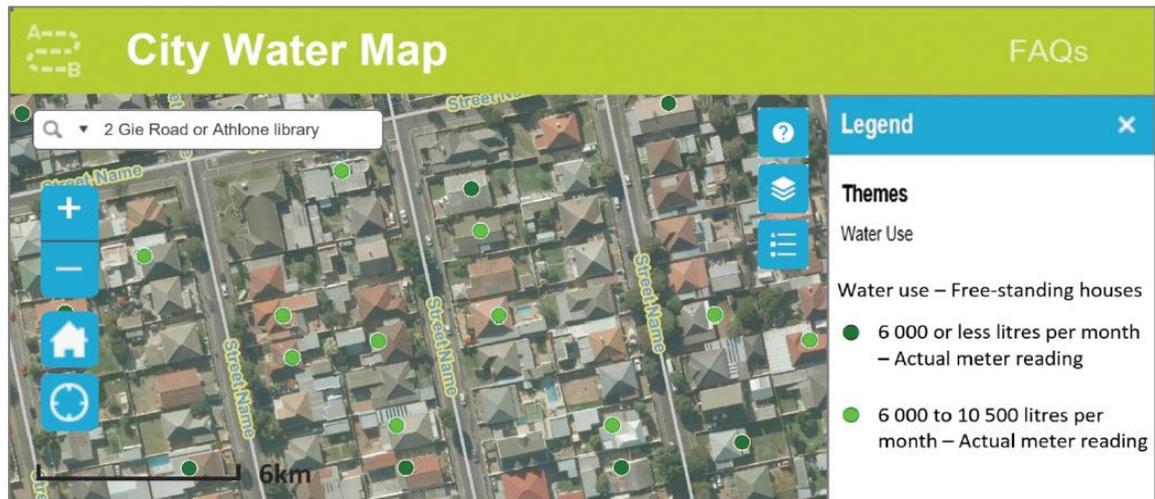
---

<sup>57</sup> William Hagman, “When are nudges acceptable? - Influences of beneficiaries, techniques, alternatives and choice architects”, *Linköping Studies in Behavioural Science* No. 213 (2018)

<sup>58</sup> Maria Bernedo, Paul J. Ferraro and Michael K. Price, “The Persistent Impacts of Norm-Based Messaging and Their Implications for Water Conservation”, *Journal of Consumer Policy* 37 (September 2014) :437-452

<sup>59</sup> Ken Sinclair-Smith, Susan Mosdell, Gisela Kaiser and Ziyaad Lalla, “City of Cape Town’s Water Map”, *American Water Works Association* 110 (September 2018): 63

**FIGURE 1** The Cape Town Water Map's online spatial viewer



FAQs—frequently asked questions

The spatial viewer uses geographic information system data of the city to display markers (light- or dark-green dots) on properties that either used less than 10,500 L (2,800 gal) or less than 6,000 L (1,600 gal) of water per month.

Figure 3

### What do the dots mean?

Household water use is indicated on the map as follows:

- **Dark green dot:** household using 6 000 or less litres per month (based on actual water meter reading).
- **Light green dot:** household using between 6 000 and 10 500 litres per month (based on an actual water meter reading).
- **Grey dot (with small dark green centre):** estimated water meter reading of less than 6 000 litres per month.
- **Grey dot (with small light green centre):** estimated water meter reading of less than 10 500 litres per month.
- **Grey dot (solid)** excluded property (sectional title property or group housing / undeveloped property / water use is zero / no available information for the property / estimated water meter reading of more than 10 500 litres per month)

**Note:** Where a property does not have a dot, it may be because:

- the property is registered on the billing system as not-residential tariff category.
- the household uses more than 10 500 litres per month (the map does not show households that use more than 10 500 litres per month).
- there could be a data issue preventing the water usage information being linked to a particular point in the map.

Figure 4

## 4.2 The Water Map: nudge or sludge?

As I have already mentioned in Chapter 1, Richard Thaler, one of the two authors of the original book about nudging, has shared three principles that should guide the use of nudges. Thaler believes that nudges ought only to be used for ‘good’, and never for evil. The three principles distantiate nudges from sludges. In this first section of the ethical analysis of the Water Map, I will take a closer look at these principles to see if the Water Map passes for a nudge, and not for a sludge.

1. *The nudge should be transparent and never misleading.*

The Water Map passes for transparency and not being misleading, as it does exactly do what it promises: it shows the water consumption per household. It is extremely transparent in the sense that it makes information that was previously not available open and visible for everyone who is interested.

2. *It should be easy to opt out of the nudge (for example, in a single click).*

It is not immediately clear whether this principle applies to the Water Map. When you look at an individual as the one who is subject to the nudge, there is a way to opt out of the nudge. This would entail to simply not go to the website of the Water Map, or to close the screen (with a single click) when on the website. This individual would not have to inform themselves with the social norms around water consumption if they did not want to. Therefore, as a subject of the nudge, it is easy to opt out of the nudge.

However, there is a second aspect of the Water Map, as residents of Cape Town were not only subject to the nudge, but also part of the nudge. Their household was on the Water Map, and therefore the Water Map partly consisted of their data. For this aspect, people were not able or allowed to opt out. They did not have a choice if they wanted to be part of the Map or not. Therefore, one could argue that it was impossible to opt out of the nudge.

3. *The behaviour being encouraged should improve the welfare of those being nudged*<sup>60</sup>

This principle does hold for the Water Map. The behaviour that was encouraged would, in the long run, definitely improve the welfare of those being nudged. If the residents of Cape Town did not dramatically decrease their water consumption, the city would risk running out of water completely, which would lead to an environmental crisis with devastating consequences. It was therefore in everyone’s advantage that people would limit their water consumption.

---

<sup>60</sup> Richard H. Thaler, “The power of nudges, for good and bad,” *The new York Times* (October 2015) <https://faculty.chicagobooth.edu/richard.thaler/assets/files/goodandbad.pdf>

## Conclusion

At the first sight, it seems the case that the Water Map classifies as a nudge, and not as a sludge. The policy intervention is transparent and not misleading, easy to opt-out of, and encouraging behaviour that will improve the welfare of those being nudged in the long run. We can therefore conclude that this policy intervention is ‘nudging for good’.

However, the second principle did make clear that there is a two-sidedness about the ability to opt-out of the nudge. Opting out as a ‘receiver’ of the nudge is not problematic, and it seems that this principle focuses especially on that. However, I have found that it was impossible to opt-out of being ‘contributor’ to the nudge, meaning that all residents of Cape Town contributed to the social norms data that was put online. This might not be problematic for Thaler’s criteria for nudges, but it does point out a problem in the design of the Water Map. The next section will go deeper into this problem, as this section focuses more on the data-sharing components of the Water Map.

### 4.3 The Water Map and the violation of privacy

I argue that the Water Map had, at the least, two implications for the privacy of residents of Cape Town. The first one is that personal information was made public by the City of Cape Town. The second one is that the Water Map placed certain individuals in a false light in the eye of the public. These implications concern two out of the four privacy rights that William Prosser claimed had arisen from different interests in informational privacy (see Chapter 2).

#### 1. Personal information was made public by the City of Cape Town

Even though the residents of Cape Town expressed their concerns about the violation of privacy, the city’s mayor stated in a news article that her intervention was not violating anyone’s privacy. She argued that the publication of data on water consumption could not violate privacy because ‘*Water consumption data relating to property is not personal information*’<sup>61</sup>.

However, I argue that the water consumption data was, in fact, personal information for Cape Town’s residents due to the circumstances. Whether something is seen as personal or private information is not only dependent on the inherent nature of the information but also because of the circumstances. At first sight, water consumption does not seem like a source of personal information. However, exactly in times where circumstances change, like the drought in Cape Town, regular information might all of a sudden become personal information. The very fact that displaying this information on an online map could be effective in changing people’s behaviour proves that it is, in fact, personal information. If this information would not tell anything about what kind of a person you

---

<sup>61</sup> Bobby Jordan, “Useful water-saving tool? Or has Cap Town gone dotty? ”, *TimesLive*, 15 January 2018. <https://www.timeslive.co.za/news/south-africa/2018-01-15-useful-water-saving-tool-or-has-cape-town-gone-dotty/>

are (a water-conscious person, or a wasteful person), the tool would not work, because people would not have felt an urge to match their behaviour to the behaviour that was the social norm. So, although the mayor (or the mayor's spokesperson) might be right that the data is usually not personal information per se, I argue that because of the circumstances during the drought, the data should have been considered personal information. Therefore, personal information was made public by the City of Cape Town, resulting in the violation of the residents' informational privacy.

## 2. The Water Map placed individuals in a false light in the public eye

The goal of the Map was to locate the households that were keeping under the daily water consumption restriction by displaying a green dot on the map. However, when a house on the map did not have a green dot, this very fact could also tell you something about the household; namely that there was a big chance that this household was using more than 10 500 litres per month. Households that were exceeding this limit were easily labelled with a negative stigma. However, the chances are that the 'label' that these households, or individuals got, was false. At the point where the drought was at its worst, the water restrictions (level 6 restrictions) meant that every individual was allowed to use 87 litres of water per day. For a four-person household, this entailed that each individual could use up to exactly their allowance of water per day, in order to stay under the limit of 10.500 litres of water per month per household, resulting in 'earning' a green dot on the map. However, households consisting of more than four individuals would not have this same fate. A household consisting of five individuals would not show up on the Water Map as a household that kept to the water restrictions even if every individual was only consuming 80% of their individual share of water. This means that individuals who were behaving exactly to the social norm and the law, were still placed in a false light in the public eye, despite their efforts.

For the reason that the Water Map made public personal information about residents, as well as for the reason that some individuals were placed in a false light in the public eye through the Water Map, I have to conclude that privacy was certainly violated. Our next question, however, focusses on the question whether this violation of privacy has any implication for the moral justification of the intervention.

### **Moral permissibility**

In the previous section, I have concluded that the Water Map did indeed violate the privacy of residents of Cape Town. In this section, I will scrutinize whether this violation of privacy leads to the moral impermissibility of the policy intervention. In order to do this, I will take a closer look at the privacy violations to see whether they were morally justified or not.

The first distinction I want to set straight is the difference between a legal permissibility and a moral permissibility. In the reactions from the municipal government of Cape Town about the Water Map, the mayor seems to focus on the legal permissibility in her defence of the intervention. In an interview, the mayor of Cape Town stated that this data was not personal and that the government was allowed to do with it what they wanted. This might, legally, be the case. However, the legality of an act or thing does not automatically lead to the moral permissibility of that same act or thing. Some legal acts can still be morally impermissible. To act morally in certain situations often requires an understanding of what is right and what is wrong, *in a certain situation*. Even though there probably are several circumstances imaginable in which sharing water usage data would not lead to any problems whatsoever, the Cape Town drought was not this type of circumstance. During the drought, water was extremely scarce, and people were told that Day Zero (the day that the town would be out of water completely) was approaching. This led to a mentality of scarcity with the residents of Cape Town. Every drop of water that was being used now could not be used later, and with the emptying dams in sight, people were almost quite literally counting the drops. This created an environment in which ‘harmless’ data on water usage all of a sudden became private, sensitive information on who was taking what in this tragedy of the commons-situation. The fear for stigmatization and harassment of properties of people that did not earn their green dot on the map became a real, and realistic fear in those circumstances. The morality of sharing this data therefore became questionable. This leads to a conflict between two values: privacy, and safety.

In the case of the Water Map, it is both important to look at the duties that the government has towards its residents, as well as the consequences that come from certain actions. One could argue that the government had two duties that are important in this case: to ensure people’s informational privacy, and to keep people safe from an environmental crisis (by ensuring their right to subsistence). In a situation where two duties are clashing, one has to question whether one of the duties can legitimately be broken. One way to do this, is to take an all-things-considered-approach. In such an approach where two duties are conflicting, one has to choose the duty that is more pressing or more important. Choosing for the more pressing or important duty over the other then leads to a legitimate reason to give precedence to that duty.

In order to answer the question whether the Water Map was morally permissible, we therefore have to assess which duty was more pressing or important: ensuring privacy, or ensuring subsistence. This leads us to a debate similar to the privacy-security debate that I have discussed before, where privacy is often put aside ‘temporarily’ to prioritise a more pressing value. This can be explained by the fact that privacy is a long-term concern, and that the ‘competing’ value is often a short-term problem. In this case that value is keeping people safe from an environmental crisis by ensuring their right to subsistence.

In this case, not choosing for subsistence could actually mean that the city of Cape Town would run out of water completely, and that people’s lives would be at stake. After all, water is of more urgent

necessity in a context of severe drought, then privacy. The all-things-considered approach therefore leads us to the conclusion that the duty to protect residents from an environmental crisis is more pressing than the duty to respect privacy.

However, this line of reasoning would only be valid in the case of the Water Map if the municipal government absolutely had to choose between the duties to safeguard privacy and to protect residents from an environmental crisis. In fact, I argue that there did not have to be a conflict of duties in the first place. The Water Map was not the *only* intervention, or even nudge, that the government could have resorted to. There are other policy interventions and nudges that have proven to be effective, but that are less controversial and are not violating privacy. Therefore, the duty of respecting privacy did not absolutely have to be violated in order to protect people from an environmental crisis. There could have been a situation where one could argue that the government had to make this choice between values, which is a situation in which research had shown that the Water Map would be the best possible intervention to change human behaviour and to decrease water consumption. This, however, is not the case. As I have demonstrated in Chapter 1, social norms nudges have proven to be able to significantly nudge behaviour in the direction of the desired behaviour, but there is no evidence that social norms nudges work better when you make private information public to everyone to see. It is clear that, for the nudge to work, people have to be informed about their own water consumption and of the commonness of the desired behaviour. Research suggests that a simple letter in the mail providing this information would have been sufficient to achieve behavioural change.

The second mechanism, however, did rely on a ‘rewarding-mechanism’, that would lead to normalization and incentivisation of water conservation behaviour. In order to achieve this goal, it seems like it was in fact beneficial and effective to reward people with a certain social status. However, this could have also been done in a different way that better takes into account privacy concerns. Alternatively, the City of Cape Town could have organized a social media campaign where people were free to share their water consumption data from the previous month. This would have led to the same feeling of being rewarded with a certain social status, and it would also have normalized and incentivised water conservation behaviour. The big difference is, that the voluntariness of this alternative would have taken away the privacy concerns that made the Water Map problematic.

For the reasons that the intervention was not strictly necessary, not proven to be the absolute most effective intervention, and the fact that there were alternatives that relied on the same effective behavioural mechanisms, I have to conclude that there was no need for the government to choose between privacy and subsistence. Therefore, the violation of privacy and the negative consequences that arose from that were avoidable and unnecessary. Because the violation of privacy and its consequences could have been avoided, but were chosen not to, I conclude that the implementation of the Water Map was morally impermissible.

#### 4.4 The Water Map and public shaming

When the Water Map was published online, residents of Cape Town expressed that they felt like the municipal government was trying to publicly shame the people who did not keep to the water restrictions<sup>62</sup>. It is debatable, however, whether the Water Map was really publicly shaming people. The Water Map, by design, displayed the people who were either using between 6.000 and 10.500 litres of water, or 6.000 litres or less. These two categorizations represented ‘good’ behaviour and ‘very good’ behaviour. The households that showed a dark- or light-green dot on the map were, therefore, being named, but not shamed. One might even argue that they were being rewarded with social recognition for their good behaviour. This can also be substantiated by the information on the website of the municipal government of Cape Town, that states:

*‘The main purpose of the map was to publicly acknowledge or “reward” households that saved water, thereby normalizing and incentivizing water conservation behaviour.’<sup>63</sup>*

The households that did not keep to the restrictions, however, were not being rewarded, as their property did not show up on the map with any kind of dot. Therefore, these households were not given any recognition for any kind of behaviour. One can argue that when someone is *not* being given any recognition for a certain type of behaviour, this can possibly imply that that person is at the same time being shamed for that behaviour.

On the other hand, one can argue that a household or person *is, in fact*, being recognized for their behaviour when their house does not show up with a green dot on the Water Map. Because of the context of the situation, the absence of ‘a dot on the map’ implies a certain kind of behaviour, which is being pointed out when they are not being publicly acknowledged or rewarded. An example will demonstrate what this situation might entail when placed in a different context. Let’s imagine a classroom full of children where the teacher praises the children who have demonstrated good behaviour by putting a sticker on their forehead. While the majority of children get a sticker, there are some children who are left without a sticker. Even though it might be the intention of the teacher to reward good behaviour, or to set a social norm of good behaviour in the classroom, the public visibility of the reward can lead to a feeling of shame by the children who did not get the reward.

The example of the teacher, the children and the stickers have demonstrated that *not* being given a visible reward can also lead to the feeling of shame for an individual. Especially when the good behaviour is being deemed with strong feelings of approval, and not participating in the good behaviour

<sup>62</sup> Olivier DW., “Cape Town’s map of water usage has residents seeing red”. *Conversat*. January 2018; <http://theconversation.com/cape-towns-map-of-waterusage-has-residents-seeing-red-90188>

<sup>63</sup> <https://use.metropolis.org/case-studies/city-of-cape-towns-water-map> > Family and Home > Residential utility services > residential water and sanitation services <https://www.capetown.gov.za/Family%20and%20home/Residential-utility-services/Residential-water-and-sanitation-services/cape-town-water-map>

is being deemed with strong feelings of disapproval, the social context can lead to a real experience of being shamed for the person who does not visibly get the approval for their behaviour.

Therefore, I conclude that in some social contexts, the not being visibly given a reward when others are, equals that person being shamed for their behaviour. Because of the design of the Water Map, this could have been the case here.

Although I have concluded that the design of the Water Map allowed for public shaming (whether it was intended or not), this does not have to mean that this public shaming lead to morally problematic consequences. Reflecting back on the three incentives for public shaming: to punish, to inform and to criticise, it becomes clear that the Water Map, above all else, attempted to inform residents of Cape Town about the norm of low water consumption. By doing so, it also informed people about who was not behaving according to the norm. However, as mentioned at the start of Chapter 3, shaming practices can often have multiple incentives and/or outcomes. In the case of the Water Map, what aimed to be an informational website, ended up being a website where people felt they were being criticised for their behaviour. The Water Map, therefore, both informed and criticised people, even though the makers of the Water Map claim that they only aimed to inform.

### **Moral permissibility**

After concluding that the Water Map was a naming and shaming policy intervention that both informed and criticised people, we now have to answer the question whether this naming and shaming was morally permissible or not. In Chapter 2 I have given three scenarios for when naming and shaming can really become morally problematic. The first scenario is when naming and shaming leads to ‘kangaroo court’ practices, where no notice is taken of standards of law or justice. This did not happen in the case of the Water Map, partially because the behaviour that was being shamed (water wastage) was already illegal under the new water restriction laws in the city. People did not take justice into their own hands to deal with water wasters (even though this was a serious concern that residents had). The second scenario, where consequences from naming and shaming become disproportionate to the conduct, did not take place either. I argue that the main reason for why did this not happen was partially because no-one could really prove that someone was wasting water intentionally. This was the case because there were multiple reasons for not ‘earning’ a green dot on your address on the Water Map. Not only the people who did exceed their allocation of water did not show on the map, but also the people that had a defect water meter, or who had a leakage in their pipeline that they were not aware of<sup>64</sup>. Therefore, there was not a one-on-one correlation between water wasters and people who did not have a green dot on the Water Map. Because these two scenarios did not take place, the third morally problematic consequence,

---

<sup>64</sup> <https://www.capetown.gov.za/Family%20and%20home/Residential-utility-services/Residential-water-and-sanitation-services/cape-town-water-map>

to induce psychological problems with people that are being shamed, did also not take place as it is usually a result of one of the first two scenarios.

A positive effect that the Water Map had, was that people really had a social incentive to start saving water, exactly because they did not want to be labeled water wasters. It can be argued that this was a positive result from the Water Map, exactly because the Water Map did allow for shaming. Fortunately, the negative consequences that could have resulted from the shaming practices did not take place because a perfect one-on-one correlation between the Water Map and actual water wasters could not be made, mainly because there were alternative circumstances that could have led to a household not 'earning' their green dot on the Water Map.

For all the reasons above, I have to conclude that the naming and shaming practices caused by the Water Map were not morally problematic, mostly because the Water Map did not allow for a perfect one-on-one identification of water wasters. Because of this, the negative consequences from naming and shaming were kept to a minimum.

#### **4.5 Conclusion**

This Chapter aimed to do an assessment of the moral permissibility of the Water Map. Section 4.2 has provided evidence that the Water Map indeed classifies as a nudge, and not as a sludge. Because of this classification, I can rule out that the intervention was morally impermissible for the reason that it was not nudging for good. Section 4.3 demonstrated that the Water Map was indeed violating the informational privacy of residents of Cape Town. I next examined whether this had implication for the moral permissibility of the Water Map. I had to conclude the value of informational privacy was being compromised for non-legitimate reasons. In Section 4.4 I concluded that, although the Water Map did allow for naming and shaming, the negative consequences from naming and shaming were kept to a minimum because a perfect one-on-one correlation between water wasters and the Water Map was not possible. Therefore, I concluded that the naming and shaming practices caused by the Water Map did not lead to a problem for the moral permissibility of the Water Map.

Taking everything that I discussed in this Chapter into consideration, I have to conclude that the Water Map was a morally impermissible behavioural intervention due to an unjustified violation of privacy.

## Chapter 5: Lessons for the future

In the previous chapters, I have spoken about the drought in Cape Town and the behavioural intervention that the city implemented to change the water consumption behaviour of the residents, the Water Map. In the ethical assessment of the Water Map, I have taken a critical look at the behavioural intervention, and I have given my judgement of the moral permissibility of this intervention. In this chapter, I will take a step back and speak more broadly about the future of nudging. I will take the lessons that we can derive from the first four chapters and use them to speak in a more normative manner about the future of nudging in public policy, social norms nudges in general, and social norms nudges in environmental crises.

The ethical assessment of Cape Town's Water Map gives rise to a larger debate about nudging in environmental crises and social norms nudging in general. As I have established in the introduction, it is time to move on clustering all kinds of nudges together, and looking at all nudges as one and the same technique. When doing this, it is easy to overlook the serious moral problems that certain nudges bring about. The broader debate about the permissibility of nudging tends to go two ways. I argue that both ways have some serious flaws and dangers. If the debate leads to the outcome that all nudges are morally impermissible, because they would be compromising autonomy or because they would be manipulative, this would include nudges that do not share these characteristics, simply because they classify as nudges. If the debate would lead to a moral permissibility of all nudges, this could potentially be dangerous because certain nudges can still prove to be morally impermissible for other reasons. The ethical assessment of the Water Map demonstrated that certain social norms nudges can cause specific ethical problems. I do not argue that the entire debate about the sometimes manipulative nature of nudges should stop. I also do not claim that nudges are never a threat to individual's autonomy. Nevertheless, I argue that there should be more debate about specific kinds of nudges, and how those nudges do or do not violate certain values.

The literature on social norms nudges demonstrates that these nudges can effectively steer human behaviour to a more desired kind of behaviour. This can be a powerful tool for governments, especially when it comes to large-scale behavioural problems, such as environmental problems. The drought in Cape Town demonstrated how collective behaviour change can safeguard people from environmental crises. However, the policy interventions also gave rise to a debate about informational privacy and naming and shaming. The specific design of the Water Map required that information about residents' water consumption was made public. This intervention was one of the first social norms interventions that used this approach. Most social norms nudges (so far) provided people in a more privacy-respecting way with information about social norms. Techniques that have been used before, as described in Chapter 1, include letters that provided some insight in statistics of the desired

behaviour. It could, of course, be interesting to test whether more privacy invading nudges such as the Water Map have more impact on steering behaviour than the more privacy-respecting social norms nudges that I have discussed in Chapter 1. If this were to be the case, this could lead to a re-evaluation of the Water Map, since the hypothetical significant impact of the intervention could justify the method. However, as this is not proven yet, it must be emphasized that we must be really careful with any policy intervention that relies on data publication.

Another point that I want to touch upon is the need for more transparency regarding the nature of certain nudges, and acknowledgement towards the existence of certain negative consequences of these nudges. When the city of Cape Town implemented the Water Map, a lot of people criticized the intervention and expressed that they had privacy concerns. The City however responded to this criticism with a statement about how the water data was legally not private data. I believe that, when there is a lot of criticism from the citizens towards a certain policy intervention, the government owes the citizens a transparent explanation of why they chose to act the way they did. The citizens have the right to know and understand why this policy intervention was put into place. Similar to when a government changes the legislation around something, for example the legal age for alcohol consumption. When a government decides to make a change in policy that impacts the citizens, the citizens benefit from knowing why this decision is made. It also often helps to create more understanding for a policy that might be infringing on people's freedom to make particular life-decisions. Overall, transparency around public policy can create more understanding and political engagement between the citizens and a government. When the municipal government of Cape Town decided to implement the Water Map, most citizens did not understand why their government chose this policy intervention over other more privacy respecting interventions, as the majority of the people do not know what social norms nudges are. More transparency about nudges could lead to a better understanding of government interventions, and more conversation between government and citizens could lead to a better understanding of the needs and concerns that citizens might have.

Another important thing that the Water Map uncovered, was the lack of acknowledgement from the government of the negative consequences that the Water Map brought about. One of the negative consequences that was raised many times was the fact that some people, and households, were put in a false light due to the Water Map. As I have already explained in the previous chapter, the Water Map was designed with the consumption limits for a four-person household. For households that consisted of more than four individuals, it was therefore close to impossible to gain the status of a household that kept to water restriction limits. The municipal government should have acknowledged and taken responsibility for this flaw in the system. The best option, in my opinion, would have been to allow (at least) these households to opt-out of the Water Map.

Another finding from the last chapter that gives rise to debate for the future of social norms nudging are the three criteria to distinguish nudges from sludges. In my assessment of the Water Map, these three criteria seemed to apply well to the nudge as a more generic behaviour steering intervention,

but it did not take into account the more specific details about *how* the nudge relied on social norms data. I found that it was impossible for individuals to opt out of being part of the nudge, although it was possible to opt-out of being exposed to the nudge. This is an important problem, because most social norms nudges rely on behavioural data of individuals. I argue that the option to opt out of being part of the nudge, especially when the data is privacy invading, is just as important as the option to opt out of exposure to the nudge when someone wants to ‘nudge for good’.

Another thing that needs to be taken into account in the future, is the evidence that we actually have for choosing one value over another. With the Water Map, it turned out that the privacy-violating design of the Water Map was never proven to be strictly necessary for the intervention to work. This is something that needs to be taken more into account, especially with otherwise evidence-based policy strategies, such as nudging.

The last point that I want to touch upon, is the responsibility that lies with the organisation that has access to certain personal data. I argue that the organisation, in this case the municipal government, has to assess whether the public disclosure of the data is a responsible thing to do or not. As I have argued in Chapter 4, the ‘privateness’ of data (and therefore also the risks for violating privacy) can also depend on the context and the circumstances and not only on the nature of the data itself. It therefore lies within the responsibility of the organisation that has access to certain data to realistically assess whether disclosure of this data is morally permissible in a certain context.

## Conclusion

In this paper, I have taken a critical look at the Water Map that the City of Cape Town implemented during the drought in the city. The Water Map was a controversial policy intervention that was not received very well by the residents of Cape Town. They expressed that it was naming and shaming people, and that violating their privacy. I have taken the concerns from the residents of Cape Town as a starting point to assess whether the policy intervention was indeed morally impermissible on grounds of these arguments.

In Chapter 1,2 and 3, I have set the foundation for my assessment of the moral permissibility of the Water Map. The first Chapter introduced the concept of nudging and the science behind social norms nudges. It became clear that social norms nudges are more prone to violating privacy and naming and shaming, because they rely on information sharing. Chapter 2 set forth the value of privacy and introduced Prosser's four privacy rights that have arisen from different interests in privacy. In Chapter 3, I have introduced the three incentives for naming and shaming and three reasons why it can become morally problematic.

In Chapter 4, I have assessed the moral permissibility of the Water Map. First of all, I looked at Thaler's criteria for 'Nudges for good'. I concluded that the Water Map did pass as such, meaning that it was a nudge that encouraged behaviour that was in the individual's best interest. Next, I determined whether the Water Map violated informational privacy of the residents of Cape Town. I had to conclude that this was indeed the case, as the Water Map violated two out of Prosser's four interests in informational privacy: 1) Personal information was made public by the City of Cape Town, and 2) The Water Map placed individuals in a false light in the public eye. After concluding that the Water Map did indeed have these implications for the violation of privacy, I looked at the moral permissibility of these violations. I had to conclude that the intervention was not strictly necessary to ensure residents' subsistence, because the privacy violating aspects about the Water Map had never proven to be necessary for the effectiveness of the behavioural intervention, and could have therefore been avoided. Because the violation of privacy and its consequences could have been avoided, but were chosen not to, I had to conclude that the implementation of the Water Map was morally impermissible.

Lastly, I looked at the Water Map and public shaming. After taking a close look at the municipal government's claim that the Water Map was not a shaming intervention, but a rewarding intervention, I still had to conclude that the Water Map did allow for public shaming. I then went on to assess whether this made the Water Map morally permissible. In order to do this, I reflected on the three consequences from naming and shaming that I argue make the technique morally problematic. I had to conclude that the Water Map did not realistically allow for any of the three negative consequences, because water wasters (and therefore the people that were most prone to being shamed) could not be perfectly identified on the map because there were different reasons for a household not to show up with a green

dot. Taking everything that I discussed in Chapter 4 into consideration, I have to conclude that the Water Map was a morally impermissible behavioural intervention due to a violation of privacy.

Chapter 5 focused on lessons for the future for social norms nudging and policy interventions that rely on information publications. I highlighted the importance of looking at every nudge as an individual policy intervention with individual implications for its moral permissibility. I also emphasized that policy interventions that rely on information publishing must be handled with extra care because of the potential moral implications. Lastly, I expressed my opinion about the need for more transparency about behavioural interventions, acknowledgement of the negative consequences, responsibility from the organisation that has access to personal data, and the option to opt-out of behavioural interventions that rely on information publishing.

The Water Map demonstrated how people can actually experience that they are living in a suffocating society (like the one Solove speaks about) when their privacy is being violated. We are living in a world where our informational privacy is often sacrificed for other values. The biggest lesson that can be learned from the Water Map is that we should, at the very least, not sacrifice our informational privacy when there is no proof that this will effectively do more good than it does harm.

## Bibliography

- Aarts & Dijksterhuis “The silence of the library: Environment, situational norm, and social behaviour.” *Journal of Personality and Social Psychology* 84(2003):18–28.
- Andrews, W. 1890. *Old Time Punishments*. New York: Dorset Press page 175
- Bernedo, M., Ferraro, P.J., & Price M.K., “The Persistent Impacts of Norm-Based Messaging and Their Implications for Water Conservation, ” *Journal of Consumer Policy* 37 (September 2014) :437-452
- Berkowitz, “Applications of social norms theory to other health and social justice issues,” In *The social norms approach to preventing school and college age substance abuse: A handbook for educators, counselors, and clinicians*. (2003), San Francisco.
- Bicchieri, Cristina, Muldoon, Ryan and Sontuoso, Alessandro, "Social Norms", *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition)
- Bloustein, E., 1964, “Privacy as an Aspect of Human Dignity: An Answer to Dean Prosser”, *New York University Law Review*, 39: 962–1007
- Cohen, *Regulating Intimacy: A New Legal Paradigm*, (2002) Princeton: Princeton University Press
- Gigerenzer, G. (2018), The bias bias in behavioral economics. *Review of Behavioral Economics*, 5 (3-4), 303-336.
- Hageman, W. “When are nudges acceptable? Influences of beneficiaries, techniques, alternatives and choice architects,” *Linköping Studies in Behavioural Science*, 213, (2018)
- Hallsworth, M., List, J.A., Metcalfe R.D., & Vlaev, I., “The Behavioralist As Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance,” *Journal of Public Economics* 148 (April 2017): 14-31

Hansen, U. & Schrader, U. *Journal of Consumer Policy* (1997) 20: 443.

<https://doi.org/10.1023/A:1006842517219>

van den Hoven, J., Blaauw, M., Pieters, W., & Warnier, M., "Privacy and Information Technology",

*The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), Edward N. Zalta (ed.),

URL = <<https://plato.stanford.edu/archives/sum2018/entries/it-privacy/>>.

Johnson, E. J., & Goldstein, D. G. (2003). Do defaults save lives? *Science*, 302, 1338-1339

Jordan, B., "Useful water-saving tool? Or has Cap Town gone dotty? ", TimesLive, 15 January

2018. <https://www.timeslive.co.za/news/south-africa/2018-01-15-useful-water-saving-tool-or-has-cape-town-gone-dotty/>

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY, US: Farrar, Straus and

Giroux.

Lehner, M., Mont, O., & Heiskanen, E. "Nudging – A promising tool for sustainable consumption

behaviour?" *Journal of Cleaner Production*, 135 (2016) : 166-177

<https://doi.org/10.1016/j.jclepro.2015.11.086>

Lourenço J.S., Ciriolo E., Almeida, S.R., and Troussard X., "Behavioural insights applied to policy":

*European Report 2016* page 31, EUR 27726 EN; doi:10.2760/903938

Marteau, T.M., Ogilvie, D., Roland, M., Suhrcke M., & Kelly M.P., "Judging nudging: can nudging

improve population health?" *BMJ* 342:228 (January 2011)

Mclaughlin, Matthew, Bellettiere John and Bliss Natasha, "One step at a time:

Simple nudges can increase lifestyle physical activity", (October 2017):

DOI10.13140/RG.2.2.32064.23042

Olivier DW., "Cape Town's map of water usage has residents seeing red". *Conversat*. January 2018;

<http://theconversation.com/cape-towns-map-of-waterusage-has-residents-seeing-red-90188>

Petley, J., "Media and Public Shaming: Drawing the Boundaries of Disclosure," Reuters Institute for

the Study of Journalism, University of Oxford.

- Prosser W., “Privacy”, *California Law Review*, 48(1960): 383–423.
- Rachels, J., 1975, “Why Privacy is Important”, *Philosophy and Public Affairs*, 4: 323–33
- Ronson, J., 1967- author. *So You've Been Publicly Shamed*. New York: Riverhead Books, A member of Penguin Group (USA), 2015.
- Rowbottom, J., “To Punish, Inform, and Criticise: The Goals of Naming and Shaming, in Media and Public Shaming”: Drawing the Boundaries of Disclosure, Reuters Institute for the Study of Journalism, University of Oxford.
- Samson & Ramani, 2018, Finding the right nudge for your clients. , *Investment News*. Retrieved from <https://www.investmentnews.com/article/20180827/BLOG09/180829939/finding-the-right-nudge-for-your-clients>.
- Sinclair-Smith et al., “City of Cape Town’s Water Map”. *American Water Works Association*, 110, (2018): 62–66
- Solove, 2008, *Understanding Privacy*, Cambridge, MA: Harvard University Press.
- Sousa et al., “The 'Day Zero' Cape Town drought and the poleward migration of moisture corridors,” *Environmental Research Letters* 13, no. 12 (December 2018) DOI: 10.1088/1748-9326/aaebc7
- Sunstein, Cass “The Ethics of Nudging,” *Yale Journal on Regulation* 32, no. 2 (2015): 415.
- Johnson et al., “Beyond nudges: Tools of a choice architecture” *Marketing Letter* 23, (June 2012): 488-501, <https://doi.org/10.1007/s11002-012-9186-1>
- Thaler, R.H., “Nudge not Sludge,” *Science* 361 (august 2018): 431-432
- Thaler R.H., & Sunstein S.R, 2008, *Nudge: Improving Decisions About Health, Wealth, and Happiness*, Revised & Expanded edition. (New York: Penguin Books).
- Thomson, Judith Jarvis, 1975, “The Right to Privacy”, *Philosophy and Public Affairs*, 4: 295–314.
- Warren, Samuel D. & Louis D. Brandeis, 1890, “The Right to Privacy”, *Harvard Law Review*, 4(5):

193–220.

Wilson A.L., Buckley, E., Buckley J.D., & Bogomolova S., “Nudging healthier food and beverage choices through salience and priming. Evidence from a systematic review, ” *Food Quality and Preference* 51, (July 2016), <https://doi.org/10.1016/j.foodqual.2016.02.009>

“Day Zero pushed back from April 16 to May 11, ” *Mail & Guardian*, February 5, 2018, <https://mg.co.za/article/2018-02-05-day-zero-pushed-back-to-may-11>

“Level 2 water restrictions imposed on Cape Town residents”, *Enca*, December 2015, <https://www.enca.com/south-africa/level-2-water-restrictions-imposed-cape-town-residents>