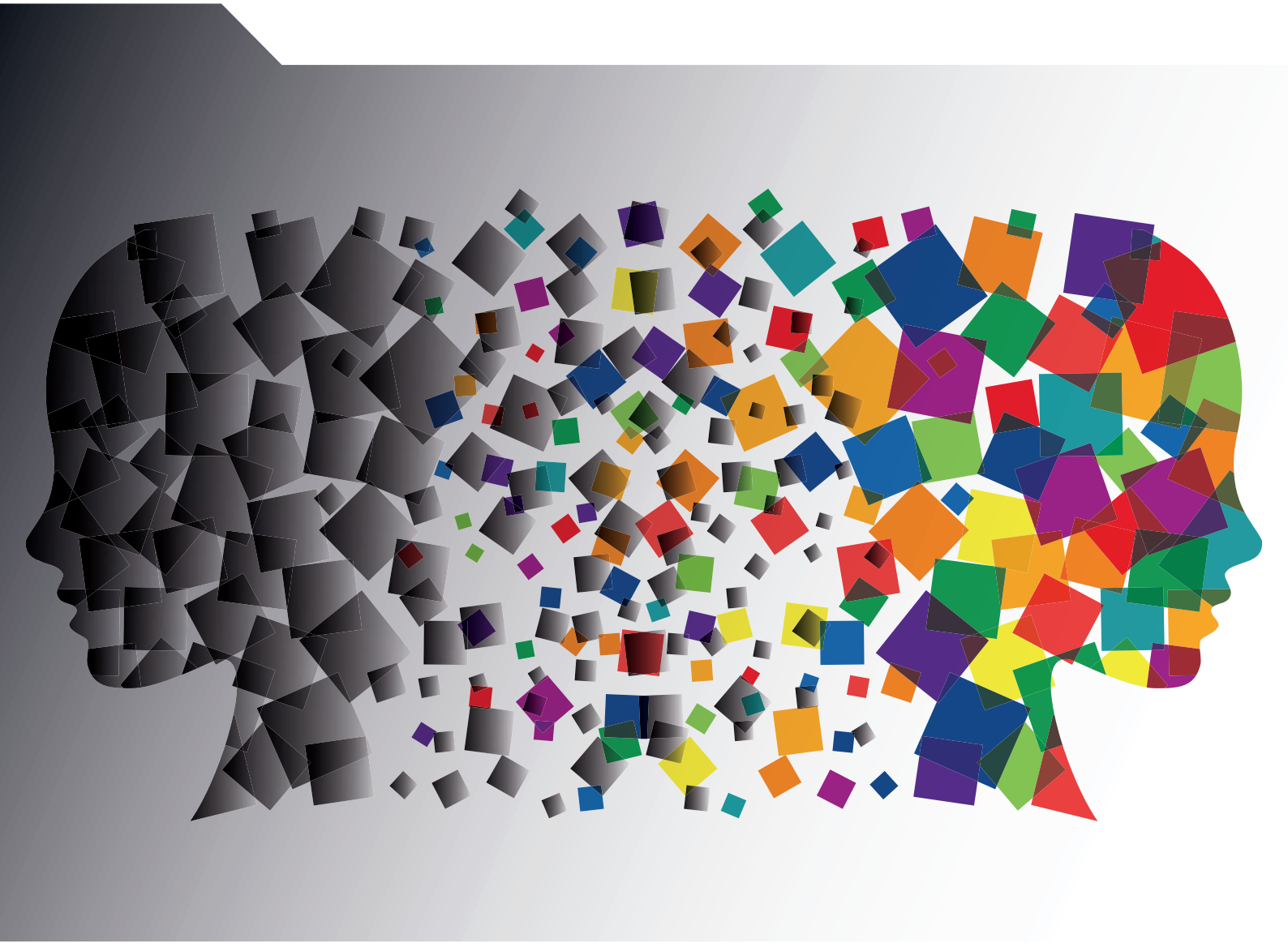
OECD Public Governance Reviews

Behavioural Insights for Public Integrity

HARNESSING THE HUMAN FACTOR TO COUNTER CORRUPTION



OECD

OECD Public Governance Reviews

Behavioural Insights for Public Integrity

HARNESSING THE HUMAN FACTOR TO COUNTER CORRUPTION

OECD
BETTER POLICIES FOR BETTER LIVES

This document, as well as any data and any map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Please cite this publication as:
OECD (2018), *Behavioural Insights for Public Integrity: Harnessing the Human Factor to Counter Corruption*, OECD Public Governance Reviews, OECD Publishing, Paris.
http://dx.doi.org/10.1787/9789264297067-en

ISBN 978-92-64-29705-0 (print)
ISBN 978-92-64-29706-7 (PDF)
ISBN 978-92-64-29729-6 (epub)

Series: OECD Public Governance Reviews
ISSN 2219-0406 (print)
ISSN 2219-0414 (online)

Photo credits: Cover © Lucky Team Studio/Shutterstock.com
© Jessica Coccimiglio/Thenounproject.com
© parkjisun/Thenounproject.com
© Sergey Demushkin/Thenounproject.com
© Gan Khoon Lay/Thenounproject.com
© Ali Coskun/Thenounproject.com
© Alvaro Cabrera/Thenounproject.com.

Corrigenda to OECD publications may be found on line at: *www.oecd.org/about/publishing/corrigenda.htm*.

© OECD 2018

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of the source and copyright owner(s) is given. All requests for public or commercial use and translation rights should be submitted to *rights@oecd.org*. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at *info@copyright.com* or the Centre francais d'exploitation du droit de copie (CFC) at *contact@cfcopies.com*.

Foreword

Integrity is more than a rational choice against corruption. Essentially, promoting integrity is encouraging behaviour in the public interest over self-serving behaviour such as corrupt and unethical practices. Yet, human behaviour is often a neglected dimension in integrity policy making. Existing efforts to preventing corruption are still widely based on a rational decision-making model. Such an approach usually stresses the importance of increasing the costs and lowering the benefits of undesired behaviour. Common policy recommendations derived from this include control and sanctions, and reducing the discretion of decision makers in order to diminish their scope for misbehaviour. Sometimes, this has led to over-regulation, the establishment of paralysing controls, and distrust in the public administration.

Integrity policies have shifted from a narrow focus on deterrence and enforcement towards promoting values-based decisions in the public sector and society. A public integrity system, as recognised by the 2017 OECD Recommendation on Public Integrity, is the foundation of trust in governments, institutions and the whole of society. With Behavioural Insights for Public Integrity, the OECD is spearheading a human-centred approach to public integrity.

Evidence from behavioural sciences highlights the social and psychological factors influencing behaviour. The 2017 Nobel Prize in economic sciences awarded to Richard Thaler is an acknowledgement for behaviourally inspired policies such as “nudges” - subtly making a good choice more likely without limiting the decision maker’s options. Over the course of the last two decades, a growing body of experimental evidence from the laboratory and the field has shed light on how corrupt networks function, how individuals are tempted to profit from corruption and how they react to the incentives provided by anti-corruption measures. Behavioural research has produced a wealth of insights that policy makers can draw from to develop innovative and well-targeted integrity policies.

This report is the first comprehensive review of different strands of behavioural sciences to identify practical lessons for integrity policies. It answers questions such as: What are the determinants of a moral choice? When are people unaware of an ethical dilemma? How is trust built and why do people lose trust in integrity? Behavioural Insights for Public Integrity guides policy makers in building integrity systems in which moral responsibility is not overlooked. It contains concrete recommendations on how to incorporate behavioural insights in modern integrity policies. At the same time, it notes aspects of existing anti-corruption strategies that could turn into behavioural pitfalls, such as the reliance on over strict enforcement or the use of transparency as a panacea. Behavioural insights can enrich integrity policies by making them not only more effective and efficient, but also more human.

Acknowledgements

This report was prepared by the Public Sector Integrity Division of the OECD Directorate for Public Governance under the leadership and guidance of Janós Bertók and Julio Bacio Terracino. It was developed by Levke Jessen-Thiesen and Frédéric Boehm with contributions from Carissa Munro, Felicitas Neuhaus, Frédéric St-Martin, James Drummond, Jennifer Eddie and Laura McDonald. Preparatory work had been done by Jovana Blagotic. Invaluable inputs and comments were provided by Christopher Swartz, United States Office of Government Ethics; Johann Graf Lambsdorff, University of Passau, and Sheheryar Banuri, University of East Anglia. The report benefitted from the editorial assistance provided by Meral Gedik, Thibaut Gigou, Andrea Uhrhammer and David McDonald as well as Edwina Collins' administrative support.

The review was approved by the OECD Working Party of Senior Public Integrity Officials (SPIO) on 15 December 2017 and declassified by the Public Governance Committee on 06 February 2018.

Table of contents

Foreword **3**

Acknowledgements **4**

Executive summary **7**

The human dimension of integrity policies 7
Central findings of behavioural sciences 7
Applying behavioural insights to integrity policies 8

Chapter 1. Introduction **9**

Chapter 2. The dynamics of moral decision making **11**

2.1. Moral choices 11
2.2. Profiting from corruption can be a temptation 14
2.3. Objectivity is an illusion 16
2.4. Justifications open up moral wiggle room 17
2.5. Corruption often feels like a victimless crime 18
2.6. Control bears a hidden cost 19
2.7. Precepts for strengthening integrity in moral choices 21

Chapter 3. Integrity in the context of social interactions **23**

3.1. Transparency is not an end in itself 23
3.2. The role of reciprocity for integrity 24
3.3. Norms and identities shape integrity 26
3.4. Enforcing norms 27
3.5. The collective action trap 28
3.6. Precepts for integrity in social interactions 29

Chapter 4. Applying behavioural insights to integrity policies **31**

4.1. A human-centred perspective on integrity systems 31
4.2. "Nudging" for integrity 48

References **52**

Tables

Table 2.1. Experimental scenario 19
Table 2.2. Main results of the experiment 20
Table 2.3. Precepts for strengthening moral choices in integrity systems 21
Table 3.1. Precepts for strengthening integrity in social interactions in integrity systems 29
Table 4.1. Use of tools for choice architecture change in an integrity system 51

Figures

Figure 1.1. Understanding behaviour for better integrity policies 10
Figure 4.1. The three pillars of the OECD Recommendation of the Council on Public Integrity: System, culture, accountability 32
Figure 4.2. Zero Rupee Note 37
Figure 4.3. Targeting a behavioural intervention 49

Boxes

Box 2.1. Moral balancing 12
Box 2.2. Australia's values alignment model 13
Box 2.3. Moral reminders 15
Box 2.4. How control can destroy intrinsic motivation for honesty 19
Box 3.1. Social preferences as justification 25
Box 3.2. Honour and identity 26
Box 3.3. The risk of extortion 27
Box 4.1. Setting standards for integrity 34
Box 4.2. The zero rupee note 37
Box 4.3. Recruitment processes and integrity – experience from Australia 39
Box 4.4. Dilemma training in the Flemish Government (Belgium) 41
Box 4.5. Guiding public officials facing ethical dilemmas in Australia 42

Follow OECD Publications on:

http://twitter.com/OECD_Pubs
http://www.facebook.com/OECDPublications
http://www.linkedin.com/groups/OECD-Publications-4645871
http://www.youtube.com/oecdilibrary
http://www.oecd.org/oecddirect/

Executive summary

The human dimension of integrity policies

Human behaviour is often a neglected dimension in integrity policy making. Traditionally, efforts to prevent corruption have been based on a rational decision-making model. This has led to policy recommendations that favour applying control and sanctions, and reducing the discretion of decision makers in order to diminish their scope for misbehaviour.

Behavioural insights enable policy makers to integrate two often-overlooked aspects into integrity policies: first, at the core of integrity is an ethical choice based on individual moral reflection. Second, social dynamics affect individual behaviour. Integrity is influenced by society, peers, neighbours and colleagues. Integrating the dynamics and pitfalls of these two core aspects can help policy makers design modern, effective integrity policies.

Central findings of behavioural sciences

Normally, policies to strengthen integrity are designed on the assumption that decision makers are motivated to act ethically. Indeed, the majority of people feel committed to integrity and think that they generally act accordingly. Yet, in reality, people are less consistent and less categorical in their ethical decisions than they admit to themselves. Sometimes a person is not even aware when their behaviour diverts from ethical standards. This is because justifications and biased judgment blur the perception of integrity breaches. Public policies can, therefore, affect individuals' moral choices by emphasising or raising their moral reference points:

- There is broad evidence that a small message, a "moral reminder", can be sufficient to induce ethical reflection. Such moments of ethical reflection can be integrated in many policies.
- Moreover, moral choices can be invoked be by creating commitments and by mentally preparing individuals for ethical temptations.
- Finally, over-strict control can have adverse effects. Excessive monitoring of a trust-based rule might drive people to disregard the rule, and create an entry-point for severe misconduct

Ethical choices are not made in isolation, but as part of social interaction. What others think or do matters. Integrity policies can be improved by taking into account the social environment in which they are applied. Guilt is smaller when shared, which means that when several people jointly engage in misconduct, each of them feels less responsible. Spreading the burden of responsibility over too many people can create an integrity risk.

Furthermore, upholding the public interest can be understood as an act of indirect reciprocity, carried out in the trust that others will do the same and the belief that integrity

benefits all. Integrity policies can reinforce such behaviour by creating an institutional setting in which the adherence to integrity is seen as the norm. Enforcing this norm, in turn, requires a balance between building trust and strictly investigating and sanctioning violations. Behavioural insights have shown that these two functions gain credibility when they are separated, e.g. executed by different institutions.

Finally, a group that has collectively de-stigmatized corrupt behaviours might find itself in a collective action trap. In this situation, an individual's motivation to act with integrity is no longer sufficient to break the vicious cycle and moral appeals will fall on deaf ears. Such situations need to be identified and systematically addressed through interventions from outside the group and strict enforcement.

Applying behavioural insights to integrity policies

This report highlights two ways in which the application of behavioural insights can lead to new, more effective integrity policies. The first is using behavioural evidence to rethink existing integrity policies and systems. Indeed, the behavioural approach addresses some of the challenges related to traditional anti-corruption policies, while emphasising the importance of promoting trust and an organisational culture of integrity. Some examples are:

- Behavioural pitfalls could be identified in the way responsibilities are assigned in an integrity system. Structures that diffuse responsibility (e.g. involving many actors, ensuring integrity merely through formal requirements) could create integrity risks.
- Ethical reflections integrated in certain procedures, especially those that carry a high risk of corruption, could ensure that integrity is not overlooked in a public official's everyday routine. Requiring a personal signature on paperwork, for example, can invoke stronger personal ethical engagement.

The second way behavioural insights can be applied is in a specific, limited intervention to fix a behavioural pitfall, improve an existing policy or encourage certain behaviour. Behavioural interventions are based on knowledge of human behavioural biases, cognitive limitations and social preferences. Such interventions often take the form of a "nudge" – a subtle change in the way a decision presents itself to help individuals make a better choice for their own benefit. Some examples are:

- Request a conflict of interest declaration from each team member at the beginning of a new project by default.
- Inform public officials that "96% of colleagues have already submitted their asset declarations" to encourage them to do the same.
- Create a quiz on the code of ethics to encourage staff to read it actively and transfer knowledge.

Behavioural sciences can provide inspiration for innovative, modern integrity policies that harness the human factor in the fight against corruption. The OECD has long supported practitioners in scoping the potential of behavioural insights, which are now being applied in various areas of public policy and public institutions globally. Exploring the implications of insights on ethical behaviour is the starting point for meaningful changes in the area of integrity and anti-corruption policies.

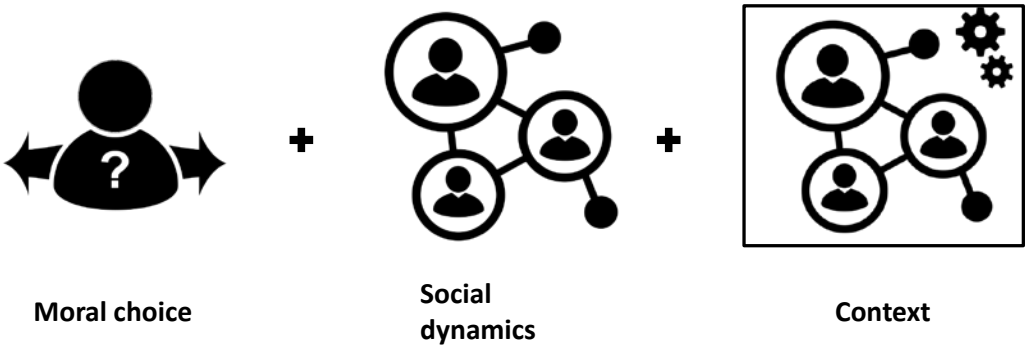
Chapter 1. Introduction

Public integrity systems aim to generate decisions by public officials that align with the public interest. As a consequence, many integrity policies seek to directly or indirectly guide and govern human choices or behaviour. In spite of this, the behavioural perspective is still neglected in integrity policy making. Researcher in various scientific fields (including psychology, neuroscience, sociology and economics) on human cognition and behaviour has produced findings that offer valuable lessons and new perspectives. As a result, policy makers worldwide are increasingly applying *behavioural insights* to other areas of public policy (Economist, 2017[1]); (European Commission, 2016[2]); (OECD, 2017[3]); (World Bank, 2015[4]).

This report presents relevant *behavioural insights* for integrity policies. A broad variety of behavioural science literature was reviewed to identify the research most relevant to integrity policy making. Indeed, rethinking integrity policies from the perspective of human behaviour can make integrity policies both more effective and efficient. Behavioural policy design provides decision makers with a different perspective that takes into account two often overlooked aspects. The first relates to the ways in which individuals make moral choices: psychological shortcuts, misperceptions and temptations can often divert the best intentions. Understanding the dynamics and pitfalls of moral choice helps policy makers provide more effective guidance to promote integrity in the public sector and the whole of society. The second aspect relates to the ways in which social dynamics impact individual behaviour. Integrity is not just a question of individual moral choice; it is influenced by society, peers, neighbours and colleagues. Integrity policies function better in congruence with – and not against – the logic of social interactions.

Integrating these two core aspects can help policy makers address some of the key challenges of modern integrity policies: guiding moral choice making, overcoming the difficulties of collective action and reducing integrity risks. Chapter 2. and Chapter 3. present core findings from behavioural ethics and social interaction research, and examine the internal process of moral decision making and the relevance of external social dynamics. Based on these insights, they identify core precepts to guide policy makers. Figure 1.1 provides a schematic overview of how behavioural insights can improve understanding of ethical decision making in a given context and thereby inform integrity policies.

Figure 1.1. Understanding behaviour for better integrity policies



To help make the practical implications of these behavioural insights more tangible and concrete, Chapter 4. applies core precepts of behavioural science to the framework provided by the OECD Recommendation of the Council on Public Integrity (OECD, 2017[5]). Finally, this chapter examines the potential application of behavioural insights to public policy within an integrity system through changes in the choice architecture – an approach otherwise known as "nudging".

Chapter 2. The dynamics of moral decision making

Integrity policies are often based on the understanding that people will exploit an opportunity for misconduct whenever its profits are worth risking potential negative consequence. In other words, individuals are assumed to weigh the probability of getting caught (i.e. the strictness of internal and external control and detection mechanisms) and the ensuing sanctions against the undue gain they could obtain through action or inaction. The policy implications of such a view involve a high level of monitoring and enforcement. However, in everyday life, morality is often perceived in terms of character traits. When confronted with an integrity-related choice, it is assumed that a person will consult his or her own personal values and act in accordance with them. Indeed, corrupt behaviour is often interpreted as an extension of the "bad character" of those involved.

Behavioural research challenges this strictly rational cost-benefit analysis as well as the limited understanding of morality as a character trait. Instead, it paints a more nuanced picture of the dynamics behind moral choices. In a well-functioning integrity system, individuals use discretionary power wisely, judge fairly, lead by example and speak up against misconduct. Reality, of course, is more challenging. How can integrity policies encourage and support such moral choices? The following sections explore the dynamics of moral decision making and examine how policies can support individuals to act with integrity.

2.1. Moral choices

Who do policy makers have in mind when designing integrity policies? Put another way, who acts unethically? The answer to this question is central to the targeting of integrity policies. A key behavioural insight is that unethical choices are often made by people who in general want to behave ethically. However, many integrity policies do not take into account the possibility that the moral compass of these individuals may fail. Some target criminal-minded actors with a view to deterring unethical choices through strict sanctions. Others rely fully on the judgement and actions of trustworthy individuals. Behavioural insights make the case for integrity policies to be based on an understanding that even well-meaning decision makers sometimes fail to make the best ethical choices, and could work to prevent these ethical shortcomings.

Understanding who commits misconduct in which kind of situation often determines the type of prevention employed. However, misconduct is not always a deliberate decision or an action based on ill intent. Rather, moral choice is a dynamic process. Evidence shows that individuals balance their moral choices by continuously comparing their current moral self-perception with their own moral reference point. The moral reference point represents the level of integrity individuals perceive as morally acceptable for themselves. If they find that their own action deviates too much from their personal reference point, they counteract (see Box 2.1). Some people tend to have relatively high moral reference points, while others might allow themselves more selfish acts.

Box 2.1. Moral balancing

"Bad" actions often produce a desire on the part of individuals to perform an act of altruism or cleansing. Conversely, doing something "good" may produce a feeling of entitlement to cheating or egoistic behaviour. This process of moral balancing is often unconscious. For example, an individual might give a donation to charity to relieve a feeling of guilt without being fully aware of this process (Sachdeva, Iliev and Medin, 2009[6]). Likewise, individuals that have acted in a particularly altruistic manner may feel licensed to be a little selfish.

Moral balancing is not limited to a specific context. The guilt acquired in one situation can make a person act in a moral fashion in a completely different one. For instance, an individual might give a larger tip at dinner because they feel guilty about stealing a colleague's dessert from the fridge at lunch.

Furthermore, moral balancing is subjective. Irrespective of the actual moral consequences of their actions, people might feel more moral due to the weight of an experience. For example, enduring pain or being treated unfairly can make people feel licensed to be egoistic (Houser, Vetter and Winter, 2012[7]; Ploner and Regner, 2013[8]). "I deserve this" is a typical justification for corrupt behaviour (see also section 2.4).

Source: (Carlsmith and Gross, 1969[9]); (Gneezy, Imas and Madarász, 2014[10]); (Kalanthroff, Aslan and Dar, 2017[11]); (Sachdeva, Iliev and Medin, 2009[6]); (Zhong and Liljenquist, 2006[12]).

The following key conclusions can be drawn from behavioural research:

- The majority of people follow a moral compass, which orients them even when no one is watching. However, they also tend to give themselves some leeway and may wander away from the moral path when given an opportunity (Mazar, Amir and Ariely, 2008[13]). This applies in particular when the immediate benefits are appealing and the costs are remote, unlikely or in the future (Gino et al., 2011[14]).
- Nobody likes to lose out. People make choices in their own favour both consciously and unconsciously and seek advantage for themselves and those similar to them.
- People do not like to be confronted with their own unethical behaviour. Instead, they may apply justifications to make the dissonance between their moral standards and their actual behaviour appear less grave (Barkan, Ayal and Ariely, 2015[15]). When they do acknowledge the dissonance, they often feel bad about their behaviour, resulting in a desire to compensate (see Box 2.1).

The silver lining for policy makers promoting integrity is that most people want to be ethical. The opposite – exhibiting no concern for the impact of one's actions on others – is a symptom of psychopathy (Hare, 2006[16]). While it remains unclear to what extent "successful psychopaths" with these attributes are present in society (Boddy, Ladyshewsky and Galvin, 2010[17]); (Stevens, Deuling and Armenakis, 2012[18]), they do not constitute the majority. The average person has an intrinsic motivation for honesty and cares about others (Aquino and Reed, 2002[19]). The strength of this intrinsic motivation depends on the individual's moral identity, prevailing social norms and other

outside factors, and varies between people and across situations and circumstances. Integrity policies should therefore target the majority and be designed to support the dynamics that encourage ethical choices. For example, Box 2.2 presents a framework used within the Australian integrity system to effectively target integrity policies to the majority of individuals who do not exhibit criminal tendencies.

Box 2.2. Australia's values alignment model

The Australian values alignment model is based on the supposition that not everyone is the same, and that not all individuals in an organisation share the same values. The following three groups are assumed to exist and represent different relative shares among the employees of any organisation:

- Group A consists of people unlikely to act in a corrupt manner regardless of circumstances, perhaps as a result of a strong moral identity.
- Group B comprises people whose decision to act in a corrupt manner is dependent on the circumstances. In ideal conditions, this group is unlikely to act in a corrupt manner. However, the opposite may be true if personal or environmental circumstances are conducive.
- Group C consists of a small group of people who tend to act in a corrupt manner whenever they can avoid the consequences. This group is driven by self-interest and tends to respond only to effective deterrence.

An effective integrity framework accepts the existence of each of these categories of people and is accordingly designed to:

- provide a work environment for staff in which high professional standards are valued, opportunities for corrupt conduct are minimised and compliance with integrity measures is facilitated (providing Group A and B with ideal conditions);
- prepare for the existence of purely self-interested individuals (Group C) by putting in place effective detection and deterrence measures.

The goal of the integrity framework is to obtain an appropriate balance of measures by encouraging integrity without losing sight of the risks.

Source: Based on information provided by the Australian Commission for Law Enforcement Integrity, www.aclei.gov.au/corruption-prevention/key-concepts/values-alignment.

Essentially, integrity policies can impact the internal dynamics of moral choice making and encourage individuals to follow their moral compass by working to influence the moral reference point of an individual in two ways:

- **Raising the moral reference point.** Raising the moral reference point induces people to compare their own actions against higher internal standards. Clearly defining ethical expectations and emphasising the relevance of integrity and the trust placed in public officials can help to adjust peoples' moral reference points upwards. Discouraging information, surveillance and distrust can, in turn, lower the moral reference point against which a person assesses his or her own behaviour.
- **Emphasising the moral reference point.** Even individuals with very high internal moral standards sometimes fail to follow them. If this happens, a dissonance arises

between a (considered) behaviour and the moral reference point. However, people often ignore such dissonances. Addressing people's personal morality and encouraging them to reflect on the ethical consequences of their actions can lead to more ethical choices.

2.2. Profiting from corruption can be a temptation

Staying true to one's own internal moral reference point is not always easy. Acting against this moral compass is often prompted by temptation – a behaviour engaged in for immediate benefit that probably does not serve long-term goals. An integrity breach sacrifices future welfare for immediate gratification (Bazerman and Tenbrunsel, 2011[20]). Actions that compromise integrity such as keeping undue travel expenses to buy a new watch, might not result in guilty feelings immediately – until the next morning when colleagues admire the watch. Integrity, in these situations, can be understood as a matter of self-control.

Time pressures and tiredness reduce self-control and have also been found to reduce ethical behaviour (Kouchaki and Smith, 2014[21]); (Shalvi, Eldar and Bereby-Meyer, 2012[22]). Similarly, intense workloads, performance pressure and short-term contracts can undermine integrity. Such conditions give public employees reasons to cut corners and circumvent rules, and may also provide them with justifications for their misconduct: "If my employers don't care about me, why should I care about them?"

In the face of an immediate temptation, people often act in a short-sighted manner and exhibit a lack of willpower. Behavioural approaches therefore try to help individuals incorporate long-term benefits into their decision making. For instance, people are psychologically better equipped to withstand a temptation when they know it is coming. Research has shown that this is also true for ethical temptations – anticipating them helps to resist them (Sheldon and Fishbach, 2015[23]). From a policy perspective, this implies an emphasis on identifying typical ethical dilemmas (e.g. in Codes of Conducts or integrity trainings). If individuals have the opportunity to consider a dilemma beforehand from a sober perspective, they are more likely to act with integrity when confronted with it directly. A similar argument can be made for the systematic identification and disclosure of potential conflicts of interest. The acknowledgement of a conflict of interest and the anticipation of a potential moral temptation reduce the risk of misconduct.

Instance of corruption may also be avoided by creating a commitment to integrity beforehand. Integrity policies could provide opportunities for people to commit to certain behaviour and refrain from others. The binding effect of such commitments will be even stronger if they are made publicly. General commitments to integrity and against corruption are usually less effective than concrete actionable statements, such as "I will not use these funds for any other than the dedicated purpose". Such actionable statements leave smaller moral wiggle room in the moment of decision making. For example, in the Philippines, vote-buying was reduced significantly after people made a promise not to sell their vote in the upcoming elections (Hicken et al., 2015[24]).

Finally, so-called moral reminders have also been shown to be an effective tool to counteract unethical behaviour. Ethical choices can be encouraged by reminding people of moral norms in the moment of decision making (Bursztyn et al., 2016[25]); (Mazar and Ariely, 2006[26]). Inconspicuous messages such as "thank you for your honesty" can have a striking impact on compliance (Pruckner and Sausgruber, 2013[27]). Such interventions have the most impact when timed carefully to just before the moment of decision making

(Gino and Mogilner, 2014[28]). Making the moral norm and self-perception salient in the decision-making moment helps morality, the long-term goal, outweigh acute temptation (see Box 2.3). However, the effect of a moral reminder may be reduced if an individual is exposed to it on a frequent basis, and at some point the message may no longer be noticed. Nonetheless, when policy makers place such messages in the proximity of decisions with integrity risks – which are not taken frequently by the same person – they can make a significant difference. Processes and procedures could therefore be assessed for the possibility of inserting timely moral reminders.

Box 2.3. Moral reminders

Small message, strong effect

One straightforward strategy to induce ethical behaviour is to remind decision makers of moral standards. Research has shown that inconspicuous messages can have a striking impact. In Austria, it is common practice to buy newspapers from a self-service box in public places. The papers have a fixed price, but there is no control of how much money customers actually place in the box when they take a paper. Researchers from the Universities of Linz and Vienna conducted a field experiment on these paper sale boxes. On some of the signs indicating the price of the paper they added, "Thank you for being honest". The amount paid for a paper increased significantly in these cases. Other studies have shown similar effects. For example, students cheated less when reminded of a moral norm, such as the Ten Commandments, just prior to taking a test.

In Mexico, the Secretaría de la Function Pública (SFP) in co-operation with the research centre CIDE (Centro de Investigación y Docencia Económicas) applied this behavioural insight to their gift registration policy, in order to enhance compliance. SFP sent out reminder emails to public employees required to register their received gifts. They randomly varied the text of the message. Five different types of reminder messages were sent:

- Legal: *It is your legal obligation to register received gifts.*
- Honesty: *We recognise your honesty as a public official. You are required to register gifts. Show your honesty.*
- Impartiality: *Receiving gifts can compromise your impartiality. When you receive a gift, register it.*
- Social: *More than 1 000 registrations per year are made by your colleagues. Do the same!*
- Sanction: *If you receive a gift and you do not inform us, someone else might. Don't get yourself punished. Register your gifts.*

The study then observed the number of gifts registered around the Christmas period (peak season for gifts), and compared this with previous years and against a control group who did not receive any of the messages. The study demonstrated that receiving a reminder email increased the number of gifts registered. However, some messages were more effective than others: reminding public officials of their legal obligations and appealing to their impartiality and honesty encouraged more people to register gifts than referring to sanctions or registrations made by colleagues.

This example shows that: (i) small behavioural nudges can increase compliance with an existing policy, and (ii) appealing to values and integrity changes behaviour more effectively than threatening sanctions.

Source: (Fellner, Sausgruber and Traxler, 2013[29]); (Pruckner and Sausgruber, 2013[27]); Santiago Steta (SFP), presentation at the meeting of Senior Public Integrity Officials (SPIO) in Paris on 26 November 2017.

2.3. Objectivity is an illusion

Objectivity and consistency in judgement are expected from public officials; unfortunately, they cannot always be expected from human decision makers. Both are often lacking in moral choices, in part due to the dynamics of moral choice making. Literature in this field highlights the ways in which human judgement is misled by unconscious bias. For instance, structures within an integrity system and throughout the public sector should secure the independence of regulatory judgement (OECD, 2017[30]; OECD, 2017[31]). Biases and the resulting discrimination most affect those decisions that rely heavily on human judgment. Indeed, the success of some integrity policies depends heavily on objective judgement, for example in areas such as conflict-of-interest management, auditing, and internal control or recruiting. The design of such policies therefore needs to take into account the biases in human decision making.

A famous experiment by (Loewenstein et al., 1993[32]) illustrates biased judgement. All participants were given the same detailed information about a car accident in which a car had hit a motorcyclist. Some were assigned the role of plaintiff (on behalf of the motorcyclist), while others acted as the defendants of the car driver. The participants were then asked how they would expect a judge to rule in this case. If their answer came close to the correct ruling they could earn money. Despite the incentive to be objective, the two groups provided significantly different interpretations of a fair settlement. Although the participants had no real interest in the case, simply adopting the hypothetical position of either defendant or plaintiff biased their judgement.

Self-serving bias can lead to favouritism not only towards oneself, but also with respect to a group. People tend to assess members of their own group more forgivingly. In an experiment, British and US American participants were presented a media article on torture and later asked what they thought. The participants were found to be more likely to judge the torture as justified when the article stated that it had been carried out by their own security services, rather than those of another country (Tarrant et al., 2012[33]). Additionally, psychological closeness can lead to lenient judgements. Some people feel closer than others, either due to direct personal knowledge or because they seem similar in some respects. Judgement therefore tends to be biased when making decisions concerning these people (Bazerman and Tenbrunsel, 2011[20]); (Chugh, Bazerman and Banaji, 2005[34]). Holding someone in high regard but meanwhile recognising an unethical behaviour of this person creates a cognitive dissonance – a condition that is psychological burdensome. The mind tends to immediately seek justifications to harmonise the conflicting observations. As a result, people are less likely to notice unethical behaviour that contradicts their beliefs about a person or group they are sympathetic towards.

In light of this research, policy makers should be aware that bias can lead to conflicts of interest that might not be obvious at first sight. Aspects that are usually desirable for

public officials, such as a strong identification with the institution, high trust and close bonds between team members, can present an integrity risk in some situations. While it is rarely feasible to isolate individual decision makers from such potential invisible conflicts of interest, risk mapping could recognise unconsciously biased decision making as an integrity risk.

Because self-serving bias is not deliberate it is not effectively addressed through harsh punishments or continuous control. However, the way in which choices are presented can prevent biased decisions. Integrity policies could ensure that at-risk decision makers are able to distance themselves from potentially biasing factors. In some settings, blind judgment is an option. When participants in the plaintiff-defendant experiment read the information before they were assigned their role as plaintiffs or defendants, they came to more convergent assessments of the case (Babcock et al., 1995[35]). The same could apply, for example, to procurement officers reviewing an offer without knowledge of the identity of the provider.

Requiring argumentation both for and against a choice before making a final decision can guide individuals towards better judgement. Asking subjects to list the weaknesses in their own argument has been shown to significantly reduce self-serving bias (Babcock, Wang and Loewenstein, 1996[36]). Where relevant, two individuals could be asked to argue for one option each and then bargain over the best choice.

2.4. Justifications open up moral wiggle room

To avoid revisiting their moral self-image, people often seek justifications for their behaviour. Justifications overcome the gap between one's moral ideal and one's actual behaviour. They function as explanations or re-interpretations of a situation that help people re-align their moral reference points with their moral self-perceptions.

The literature provides various examples of typical justification mechanisms:[1]

- **Linguistic euphemisms**: Using benign labels for corrupt acts lowers the burden of ethical responsibility. For example, "cookie jar reserves" refer to the use of surpluses from profitable years to improve the balance sheet during leaner years (Economist, 2010[37]).
- **Metaphor of the ledger**: People re-interpret benefits from an integrity breach (e.g. embezzlement) as a well-deserved reward for their good efforts, in particular when they feel under-appreciated.
- **Denial of victim, injury or responsibility**: People engaging in unethical acts do not acknowledge the harm done or do not see it as a direct result of their own action. This enables them to avoid experiencing the full moral responsibility.
- **Legality/Rule-bending/Loop-holing**: When an act is unethical but legal or allowed, the responsibility is shifted to those making the law or rules. Conversely, a rule that seems unfair might be circumvented with less moral compunction.
- **Cash substitution**: Stealing or receiving non-monetary goods (e.g. taking resources from the work place) can seem less unethical than stealing or receiving money as a bribe or cheating for a monetary amount.

[1] (Ashforth and Anand, 2003[119]); (Barkan, Ayal and Ariely, 2015[15]); (Bazerman and Tenbrunsel, 2011[20]); (Bicchieri and Xiao, 2009[120]); (Erat, 2013[118]); (Falk and Szech, 2013[41]); (Hamman, Loewenstein and Weber, 2010[42]); (Kouchaki et al., 2013[114]); (Lambsdorff, 2012[55]); (Paharia et al., 2009[117]); (Pittarello et al., 2015[116]); (Tenbrunsel, 1998[115]).

- **Reframing an ethical decision as an economic one**: This mechanism shifts the focus away from the moral damage to the measurable benefits.
- **"Others do it"**: The behaviour of a direct peer is automatically perceived as more acceptable. Use of this mechanism reinforces misconduct within a group.
- **"Everybody does it"**: When a certain integrity breach is tempting, people convince themselves that many others engage in this behaviour, which makes it more acceptable for them to do so.
- **Delegation**: Superiors delegate "dirty tasks" or ask for results that are not feasible without breaching integrity, thereby relinquishing personal responsibility for the actions of their subordinates. The subordinates "do as they are told", thereby neglecting personal responsibility for the moral consequences of their behaviour.

All justification mechanisms ease the burden of immoral behaviour for decision makers, allowing them to make unethical choices in front of others, while avoiding personal feelings of guilt. They create moral "wiggle room" (Dana, Weber and Kuang, 2007[38]) – situations in which objective moral assessment can be avoided. Evidence suggests that dishonest behaviour depends on the extent to which such justifications are (Dana, Weber and Kuang, 2007[38]); (Pillutla and Murnighan, 1995[39]) and (Shalvi et al., 2011[40]). Individuals will only cheat if there is sufficient moral wiggle room to justify their behaviour to themselves and others.

This finding has a key implication for integrity policy makers: reducing moral "wiggle room" lessens the tendency to take immoral decisions. Public sector employers can contribute in this regard by defining clear and feasible expectations of proper conduct for their employers. Tasks and incentives should be aligned with the organisational goals and not contradict or undermine integrity. Employees who are well aware of the purpose of their tasks and the consequences of their actions will also have less moral "wiggle room" to justify unethical behaviour.

2.5. Corruption often feels like a victimless crime

Experimental evidence shows that the perceived (im)morality of an action decreases with the distance between the action and the harm caused by that action. Indeed, indirect agency (e.g. when an unethical task is delegated) or uncertainty regarding whether one's own actions are pivotal for the outcome reduces the sense of moral responsibility (Falk and Szech, 2013[41]); (Hamman, Loewenstein and Weber, 2010[42]).

Corrupt transactions often involve intermediaries, clandestine causalities and invisible damages. The harm done by corruption is often far removed from the corrupt actor. This makes it easier to deny responsibility and reduces the moral burden of engaging in corruption. In addition, guilt becomes smaller when shared. When the responsibility for an immoral decision is borne by several actors, the actions of one individual seem less pivotal. Common reflections in such situations include "My action doesn't make a difference" or "If I don't do it, someone else will".

Accordingly, a central consideration in the design of integrity policies could be to reduce the distance between action and harm. The design of organisational structures, communications, stakeholder engagement and human resource management in the public sector can evoke a sense of responsibility for integrity. For example, the distance between action and harm can be reduced by speaking about integrity as a personal matter rather than in abstract terms. The clearest assignment of responsibilities in the case of many integrity policies would be to hold a single actor responsible for choices. A second person

can play a controlling function in decision making, but should not share the moral responsibility. Communications focusing on outcomes and dialogue with the affected citizens can also increase a sense of pivotality among public officials – that is, the perception that their choices make a difference.

2.6. Control bears a hidden cost

The introduction of a control mechanism is a signal of distrust. Employers who rigidly monitor the actions of their staff are sending a signal to their employees that they expect misconduct. In such situations, the risks of detection might deter those with bad intentions from engaging in misconduct. However, the evidence highlights the adverse impact on intrinsic motivations for integrity (Schulze and Frank, 2003[43]): emphasising controls and sanctions may give or reinforce the impression that misconduct is common and thus more acceptable. Box 2.4 describes an experimental set-up that showed the deterrent effects of control.

Box 2.4. How control can destroy intrinsic motivation for honesty

Schulze and Frank (2003[43]) undertook an experiment to investigate whether the introduction of control reduces honesty. The experiment was carried out in the context of a movie club at a university. Just before the movie was shown, the following scenario was presented to the students:

A 200 Deutsche Mark (DM) bill belonging to the club fell into a pipe and cannot be removed without the help of a plumber. The student commissioned by the club must select the best offer from among 10 plumbing companies. The offers consist of: (i) the price the movie club will have to pay, and (ii) a bribe for the student in exchange for choosing this company.

After the film, a sheet was selected at random and the chosen student anonymously received a payment in accordance with the choice he or she had made.

The researchers organised two versions of the experiment – one with controls and one without. When controls were in place there was a chance of detection, which increased with the amount of the bribe. If the student was not detected, he or she would receive the payment according to the choice made. If the student was detected, the cheapest company won the contract and the student would receive nothing. Table 2.1 provides an overview of the situation.

Table 2.1. Experimental scenario

Company	Firms' bids of price and bribe		Probability that the student is detected	Expected value of the bribe (probability of detection x bribe)
	Price which the movie club has to pay (DM)	Amount which the student could receive (bribe in DM)		
A1	20	0	0	0
A2	40	16	0	16
A3	60	32	1/6	26.7
A4	80	48	1/6	40
A5	100	64	1/3	42.7
A6	120	80	1/3	53.3
A7	140	96	1/2	48
A8	160	112	1/2	56
A9	180	128	2/3	42.7
A10	200	144	2/3	48

The results were astonishing. The average amount of bribes was somewhat higher in the risk version then in the non-risk version of the experiment. On average, students behaved more corruptly with a risk of detection than without. In the non-risk version, 9.4% of the students were completely honest, choosing the most efficient company without asking for a bribe in return, compared to only 0.9% in the risk treatment. Control had reduced honesty. In other words, the intrinsic motivation for honesty had been crowded out. This result clearly contradicts any simple recipe for controlling corruption through stricter law enforcement. However, deterrence works too. The risk of detection had reduced the percentage of particularly ruthless students who selected the maximum bribe. This share fell from 28.8% in the non-risk treatment to 12.6% in the risk treatment (Table 2.2).

Table 2.2. Main results of the experiment

	No bribe asked (Alternative A1)	Highest bribe asked (Alternative A10)	Average bribe
risk treatment	9.40%	12.60%	91.4
non-risk treatment	0.90%	28.80%	87.9

In conclusion, the experiment showed that the addition of controls destroys the intrinsic motivation for honesty, but also acts as a deterrent to corruption.

Source: Schulze and Frank (2003[43]).

Consideration of detection risks in an ethical decision implicitly turns an ethical question into a mere cost-benefit consideration. If the expected benefit from the unethical action outweighs the expected cost, taking into account the probability of detection and the sanction, then the individual will likely choose this option. Experimental evidence suggests that such a *reframing* (see section 2.4) shifts people's focus away from the moral implications, making corruption more likely.

Furthermore, control has been shown to significantly reduce the efforts of the person being controlled (Falk and Kosfeld, 2006[44]). It forces people to provide only the minimum effort necessary to pass the control, but removes the element of positive reciprocity: employees subject to controls might feel less obliged to remain faithful, even when they are being treated well overall (Lambsdorff, 2015[45]).

Lastly, while limiting discretion in decision making is a standard recipe of traditional anti-corruption measures, limiting a person's scope of action through controls and reporting regulations might be counterproductive. Public officials with strictly reduced discretionary power, for example, have less freedom to draw on their expertise and experience to decide in the public interest.

However, the fact that controls have an associated cost does not mean that they should be removed. The detection of misconduct and the enforcement of legal norms are essential for building and maintaining trust in integrity. The policies that could be reconsidered in the light of this cost are those that induce a strong burden of distrust while offering a limited chance of detection. When compliance with a rule is generally high and a full control is less feasible, trust might be a better choice than control. Vicious control of a trust-based rule might even drive people to clandestinely disregard the rule, which creates an entry point for severe misconduct. For example, sanctioning an employee for

accepting a box of chocolates, rather than trusting that the gift will not unduly influence them, might create resentment. If employees feel distrusted, they might begin to secretly accept small gifts, knowing that such policies are difficult to monitor, and thereby develop a tolerance of non-compliance with the gift policy.

2.7. Precepts for strengthening integrity in moral choices

Individuals make all kinds of unethical choices and in many cases may not even be aware of doing so. Consideration of the main dynamics impacting moral choices allows public policies to be rethought in a way that supports moral decision making and strengthens integrity. The ways in which behavioural insights on moral choices can enrich integrity policies are summarised in Table 2.3.

Table 2.3. Precepts for strengthening moral choices in integrity systems

Precepts for strengthening moral choices in integrity systems	See section
a) Expect ethical failures from anyone b) Raise moral reference points c) Emphasise moral reference points	2.1
d) Provide actionable training and commitment e) Encourage ethical reflection (e.g. through moral reminders)	2.2
f) Anticipate biased assessment	2.3
g) Counteract justifications	2.4
h) Invoke individual responsibility i) Reduce the distance between action and harm	2.5
j) Avoid framing ethical questions as matters of control k) Avoid communicating distrust	2.7

"We are not as ethical as we think we are" is one conclusion from research into behavioural ethics. Policies that aim to strengthen moral choice making could therefore target the majority of public officials who are motivated to act ethically, yet sometimes go astray. The fact that moral choice is dynamic and can be impacted by external factors also means that policies can intervene.

According to the theory of moral balancing, policies can affect moral choice making in two ways – by emphasising moral reference points and raising moral reference points. While the former involves building long-term norms and strengthening integrity values, the latter is concerned with on-the-spot advice and reminders. Chapter 4. takes a closer look at how both approaches can be integrated into an integrity system in a behaviourally informed way.

Why do people deviate from their own moral standards? Integrity can be seen as a matter of self-control. As such it can be strengthened by creating pre-commitment and mentally preparing people for ethical temptations. Furthermore, there is strong evidence that moral reminders affect ethical choices. A small message can be sufficient to induce ethical reflection – a quick mental check of an intended action against one's own moral reference point. Moments of ethical reflection can be integrated in many policies.

Sometimes, people are not even aware that their behaviour diverts from ethical standards. Typical justifications and biased judgment can hide integrity breaches in personal choices or the local environment. Behaviourally informed policy design can help erect lines of defence against such conduct.

Lastly, the research in behavioural ethics shows that overly strict controls can have adverse effects. A careful approach to control policies that avoids communicating prevent the discouraging effects of control without abandoning the use of strict and strong measures to enforce integrity.

Chapter 3. Integrity in the context of social interactions

Moral choices are not taken in isolation. Human decisions are often driven by social motives such as loyalty, trust building, returning favours or helping someone out of a tricky situation. People take decisions in their own best interest, but they also care what others think or do. Social motives can work in favour of integrity, but they can also function as reasons for corrupt decisions. Understanding why humans as social decision makers are swayed by their own preferences, as well as those of others, can provide insights into why and how integrity policies work or fail.

3.1. Transparency is not an end in itself

The moral wiggle room for a decision maker reduces when other actors are involved. An individual when alone will be the first to be convinced of the morality of his or her actions. Outside observers who are not biased might not be as easily convinced by the justifications offered. The opinion of outside observers usually matters to a decision maker: people prefer to act in a self-serving manner, but at the same time like to appear moral to others (Batson et al., 1999[46]). Transparency and accountability mechanisms could thus reduce unethical behaviour. The perception that one's behaviour is visible, and potentially observed, introduces an element of accountability that makes it more difficult to apply justifications (see section 2.4), because potential observers could easily detect an excuse.

However, in the absence of interaction with an observer, such effects might dissipate. Transparency without interaction can be compared to a security camera without anyone to watch the tapes: those monitored may forget its existence and behave as if they were alone. Many integrity policies introduce transparency, for example, through open government and open data strategies; however, transparency is not an end in itself. From a behavioural perspective, approaches to promote transparency can have a significant impact if they trigger dialogue. Not all elements need to be subject to regular interaction, but regular attention to specific elements can increase the sense of social observability for all transparent actions. For example, a committee whose work is publicly observable and who, every once in a while, receive comments or complaints from citizens, might feel more liable to the public. The regular reminder that their decisions affect citizens reduces the perceived distance between action and harm, and thus limits moral wiggle room. This effect is not achieved by transparency, but rather through actively engaging with the respective stakeholders. Transparency is thus a necessary condition to achieve this desirable effect, but is not sufficient in and of itself.

The interaction between transparency and integrity is not always straightforward. In experiments, people who disclosed a conflict of interest afterwards gave biased advice in a more brazenly manner (Cain, Loewenstein and Moore, 2011[47]); (Cain, Loewenstein and Moore, 2005[48]); (Loewenstein, Sah and Cain, 2012[49]). This finding reveals a diffusion of moral responsibility in transparency: the person who disclosed the information feels that they can act upon their bias once it is openly admitted. They feel

less responsibility to balance their decisions, based on reflections such as "I didn't hide anything" and "If this is a problem, someone will tell me". Transparency may thus provide absolution for an unethical choice. Who granted the absolution? The other participants in the experiments usually did not expect biased advice. On the contrary, they mistook the disclosure of information as a sign of trustworthiness (Sah, Loewenstein and Cain, 2013[50]). In reality, those affected by the decision might not even be aware of the conflict of interest. In scenarios where significant volumes of information are disclosed under open government and open data policies, the onus to filter relevant information and discover a potential conflict of interest falls on the media and civil society. If a conflict of interest is hidden in disclosed information but remains undetected, responsibility may become lost somewhere between the disclosure and the reception of the information. In such cases, untargeted transparency without a clear recipient might even reduce accountability for integrity.

3.2. The role of reciprocity for integrity

The social determinants of ethical behaviour can be widely understood with one concept: reciprocity. Reciprocity is a social behaviour that can be observed – to different degrees – in every culture of the world, and is fundamental to the human species (Bowles and Gintis, 2011[51]); (Camerer, 2003[52]); (Henrich et al., 2005[53]). Someone who receives kindness will return the favour. If treated unfairly, people seek revenge. This simple reasoning lies at the core of human logic and constitutes the foundation of market exchange, trust and fairness (Fehr and Gächter, 2000[54]).

However, reciprocity also lies at the core of many corrupt interactions. Since corruption is clandestine, a corrupt deal cannot be enforced through formal institutions such as those provided by the justice system. To some degree, the actors have to trust that their counterparts will follow through as expected (Lambsdorff, 2012[55]). The persistence of human reciprocity allows them to trust that they will obtain the services that they paid for with bribes; there is such a thing as "honour among thieves" (Husted, 1994[56]).

Essentially, reciprocity can function as a motivation or excuse for engaging in corruption. Box 3.1 presents three settings where reciprocal norms might make corruption seem less immoral. For example, involving several actors in a decision can lead to a diffusion of responsibility. Therefore, the involvement of additional actors in decisions (e.g. a requirement of approval by at least two equally responsible individuals known as the four-eyes principle) might not necessarily be an effective approach to promoting integrity.

Box 3.1. Social preferences as justification

Hiding behind good intentions can deter people from admitting the actual moral implications of misconduct to themselves or others. Typical justification patterns involving social preferences include:

- **Self-serving altruism**: When someone else also benefits from misconduct, the other person's interest is used to justify the action in place of one's own. The benevolent motive, for example, helping a friend to get a job, overshadows the corrupt act, in this case, nepotism (Barkan, Ayal and Ariely, 2015[15]); (Shalvi et al., 2015[57]).
- **Robin Hood logic**: A harm done to a stronger/powerful/richer entity is justified on the basis of a preference for equality.
- **Diffusion of responsibility:** When several people engage in misconduct, the chances of one individual speaking up against it are reduced. Each individual feels less responsibility for the action and does not want to limit the other persons' freedom of choice or indicate their distrust (Moore and Gino, 2013[58]; Wiltermuth, 2011[59]; Schikora, 2011[60]).

Acting with integrity often means refraining from reciprocity – recruiters are expected to favour strangers over people they know personally based on merit, bidders should make their case with good proposals rather than gifts, and parliamentarians and judges should make impartial choices. Yet, in some settings, the patterns of reciprocity are more reliable than institutions. For example, a significant career step might not be possible without good personal connections or a public service might not be obtainable without a bribe. In such cases, it is rational to build personal connections and perform favours to expect favours in return. The fact that reciprocity is a strong human preference can in many cases explain why corruption occurs or why an integrity policy is not effective.

Strengthening integrity of course does not require eliminating reciprocity. On the contrary, it is in the interest of integrity to strengthen reciprocity, in the form of a more sophisticated *indirect* reciprocity. Instead of "you scratch my back and I'll scratch yours", indirect reciprocity follows the logic of "I'll scratch your back and someone else will scratch mine" (Nowak and Sigmund, 2005[61]). However, such indirect reciprocity has to be supported by shared norms and formal institutions – a system of laws and regulations. Unlike a corrupt interaction, the reciprocity of norms and institutions is not bilateral. It takes place between people who do not necessarily have a personal relationship. Individuals agree to serve the public interest, trusting that the public interest also serves them. Truly moral behaviour does not pay any attention to personal benefits and is solely true to the norm. "Doing the right thing" without expecting praise signals to other people that one is trustworthy and believes in the shared norm (Yoeli et al., 2013[62]). Integrity thus means remaining truthful to the public interest in the face of opportunities to abuse entrusted power for personal gain.

Making choices in the public interest requires trust in the system that it supports: for example, trust in a recruitment system and certifying institutions, trust in a fair bidding process, and a belief in justice and fairness. Eventually, integrity also means trusting that everyone in society benefits from upholding certain norms, value and rules above their direct personal interest. An individual's contribution to integrity is thus indirectly

rewarded. Strengthening trust in formal institutions and ethical norms and entrenching values in professional identities essentially strengthens indirect reciprocity.

3.3. Norms and identities shape integrity

People orient their actions in accordance with their perception of acceptable behaviour within their social context – whether this is a culture, society or peer group. Famous psychological studies such as Asch's conformity experiment, the Stanford Prison Experiment or Milgram's obedience experiment have demonstrated the extent to which people's choices are impacted by the norms inflicted and the roles prescribed to them (Asch, 1952[63]); (Milgram, 1963[64]); (Zimbardo and Cross, 1971[65]). The Stanford Prison Experiment, for example, transformed students into cruel prison guards or submissive detainees simply through external prescription of these roles by the researchers. In all three experiments the social environment induced participants to engage in cruelty and untruthfulness to an extent that surprised them as much as outside observers.

More recent evidence from research into behavioural ethics confirms the relevance of social norms and identities for moral choices. If one person lies or cheats without facing consequences, this behaviour can spread among friends or colleagues (Gino and Bazerman, 2009[66]); (Robert and Arnab, 2013[67]). In particular, gradual divergences from an ethical norm tend to be more accepted by others, creating a slippery slope towards generalised dishonesty (Gino and Bazerman, 2009[66]). A key factor in creating and strengthening behavioural norms is social identity (Akerlof and Kranton, 2011[68]); (Tyler, 1999[69]), defined as the role individuals assign themselves in a group and the group in society.

The key outcome of such research for integrity policy makers is the importance of continuously strengthening and renewing their efforts to establish integrity as a social norm and to integrate integrity into professional identities. Policies shape norms and identities in many ways. A code of ethics, for example, can be used to emphasise integrity as a common norm of behaviour. Box 3.2 presents an experiment that showed how reference to a moral norm as part of social identity increased honesty.

Box 3.2. Honour and identity

Would students perform better in a test if supplied with the answers to all questions? Researchers gave a group of students a Maths test together with the answers. On average, the students did better than a control group who took the test without knowing the answers. Evidently, the first group cheated and looked at the answers to improve their results.

Some test sheets, however, included the following sentence: "I understand that this short survey falls under [the university's] honour system". Students who received such a sheet were significantly less likely to cheat on the test. The honour system motivated them to behave more ethically. However, no honour system existed at the time at the schools where the experiment was conducted (MIT and Yale). Regardless, the effect of mentioning an honour system was the same as at Princeton University – where an honour code actually exists. The effect thus cannot be attributed to the honour system in itself, but rather to the reminder of a moral norm. The reference to honour connected with the social identity of the prestigious schools induced greater honesty in the students.

Source: (Mazar, Amir and Ariely, 2008[13]).

Integrity policies can acquire leverage where they acknowledge the power of social norms. From shaping the public debate to changing an organisational culture, social norms can play a key role in enhancing or hindering the effectiveness of integrity policies. Section 3.4 explores possible actions to strengthen integrity norms.

3.4. Enforcing norms

Behavioural research points to the conclusion that informal social norms cannot be strengthened through formal enforcement. Rewarding co-operative behaviour is often more effective than punishment in motivating co-cooperation (Rand et al., 2009[70]). Reducing an individual's freedom to devise their own rules might even hinder the development of social norms and the collective resolution of social dilemmas (Ostrom, 2000[71]). As discussed above, an overly strict regime of enforcement and control can crowd out the intrinsic motivation for integrity. While it might be possible to achieve compliance by using fear, more beneficial results can likely be achieved by encouraging public officials to follow their intrinsic motivation towards integrity. Signalling distrust reduces the moral reference point (see section 2.6) as well as trust in the shared norm.

Furthermore, a fear-based enforcement approach could eventually be misused for extortion. Box 3.3 provides an example of how fear of punishment could reinforce corrupt relationships. In some cases, both the abuse of office and the action of receiving a bribe are punishable, leading to situations where the briber may threaten to report the recipient of the bribe if they do not receive an expected favour in return. An open organisational culture, anonymous reporting mechanisms and lenient punishment of self-reporters can help prevent such situations (Buccirossi and Spagnolo, 2006[72]); (Lambsdorff and Nell, 2007[73]). However, institutional oversight, whistleblower laws and prosecution need to be designed specifically to prevent their misuse and increase their effectiveness.

Box 3.3. The risk of extortion

In 1987, an inexperienced employee of the road construction authority in the city of Bochum, Germany, accidentally transmitted the names of competing firms in a public tender. Following this incident, he received an envelope containing 2 000 Deutsche Mark from the private firm that obtained the information. Out of fear of being reported he unduly favoured this firm in various tenders in the years to come. In court, he gave the following statement justifying his behaviour and explaining how he became entrapped in this corrupt relationship: "Suddenly I knew that I had begun to be at his mercy".

Source: Lambsdorff (2002); "Mit jedem Gefallen tiefer in den Sumpf", Westdeutsche Allgemeine Zeitung (1998).

At the same time, where formal norms (e.g. laws and regulations) are concerned, enforcement can serve to build trust. Enforcing sanctions for violation of a formal norm increases confidence that others will not violate the norm, even in cases where a person's commitment to integrity raises doubts. Sanctions enable a group to trust that its members will uphold an agreed standard of behaviour, even when individual members might not have confidence in the ability of all to adhere to the agreed norms. Where misconduct is common and clearly in breach of an integrity norm, control and enforcement can be used

as tools to restore trust in the norm. Investigations of integrity breaches should therefore be thorough, and any sanctions imposed must be credible and visibly enforced in an impartial and timely manner.

At no point does trust render enforcement redundant. Even in cases where most people are committed to a certain formal norm and would still adhere to it in the absence of punishment, enforcement has a relevant behavioural function – it serves as negative indirect reciprocity. The wish to reward commitment to a norm by others (indirect reciprocity) is matched by the desire to penalise violation of a norm by others (negative indirect reciprocity). If all but one person in a group acted with integrity in the absence of sanctions, the group would still feel that this person should not profit from corruption. Accordingly, the other group members would urge the authorities to impose sanctions on the person violating the norm. The punishment of norm violations also functions as a valuation of the norm and the contributions that everyone makes to ensure it is upheld. Strict and consequent enforcement of integrity laws and regulations is therefore a tribute to integrity and all those committed to it.

Not all norms are enforceable. As argued above, efforts to enforce and control behaviours that cannot be controlled promotes distrust and crowds out the intrinsic motivation for integrity. Introducing the necessary controls separately from motivating and norm-building policies avoids undermining the intrinsic motivation for integrity (Lambsdorff, 2012[55]).

3.5. The collective action trap

A group of individuals may find itself caught in a low-integrity trap, which can be understood as a collective action problem (Persson, Rothstein and Teorell, 2013[74]). The shared norm of indirect reciprocity is what economists call a public good: if everyone contributes, then everyone benefits. Where everybody cheats, however, those who do not are simply foolish. They might even be punished by others in the group. Cheating then may become accepted and sometimes even expected. It might still be recognised as an integrity breach, but there is no severe (socially enforced) moral stigma. Take, for example, an organisation where nearly everybody over-claims travel expenses. Claiming the true cost of travel might lead a person to feel they were missing out. At the same time, the claimant would risk the disapproval of colleagues, as a low expenses claim might expose the wrongdoing of others.

In cases where systems are failing and loyalty to an integrity norm no longer benefits those who adhere to it, more direct forms of reciprocity gain relevance again. People look after the interests of their peers and clientele and expect assistance only from those whom they help. Sacrificing direct reciprocity for the sake of integrity under these conditions seems unwise. For example, if an individual takes a test, knowing that all the other candidates are secretly helping each other, not cheating will put him or her at a disadvantage when the results are compared. Under these circumstances, the individual might feel less guilty about helping their neighbour. Similarly, an individual will feel less responsible for an integrity breach in a collective action trap. The individuals in such situations may even feel that they had little choice in the matter. A true sense of guilt becomes less likely at this point, while justification mechanisms are probably already present. An integrity breach thus becomes understandable, and people can more or less openly admit their participation. Such collective action traps explain in part why corrupt actors often have no sense of wrongdoing or why a group may be surprised that a

behaviour that had become common practice among them is perceived to be wrong outside their team or organisation.

Individual organisations or services can all fall prey to collective action traps. Integrity policy makers therefore need to be aware of this risk, its underlying dynamics and ways to break the vicious circle. Appeals for more integrity can easily fall on deaf ears when they emphasise a norm that nobody sees applied in their daily lives. Policies that build on the shamefulness of integrity breaches (e.g. gift registry, asset declarations, citizen accountability through transparency) might lose their impact when a group commonly accepts misconduct. To avoid spending resources ineffectively, integrity policy makers need to know their target group well. They need to identify which norms are upheld, which behaviours are accepted and which integrity breaches are tolerated. Citizen or staff surveys can be used to gain a better understanding of the dynamics involved. Once a collective action trap has been diagnosed, targeted measures can be applied.

The vicious circle of the collective action trap needs to be disrupted by sanctioning misconduct in a *visibly* and *timely* manner and by fostering the collective action of those who are willing to do a good job, for instance by facilitating communication between them and making them visible. In this way, the vicious circle can gradually revert to a virtuous circle. Escaping a collective action trap using resources inside the system can be a tedious process, however. In extreme cases, intervention from outside or a complete change of leadership may be necessary to achieve the required shift. Usually, some type of outside support will be needed to strengthen change actors within the organisation.

3.6. Precepts for integrity in social interactions

Integrity can be defined as an individual's choice to stand up for the public interest. Yet, ethical choices are not made in isolation – they result from social interaction. Integrity shifts in line with the dynamics of the moral identity of an individual, as well as with beliefs about what others think or do. Integrity policies can therefore benefit from a design approach that takes into account the social environment in which they will be applied. Which social norms are prevalent? Which motives are relevant to the people? Which relationships are impacted? These factors and their influence on integrity should not be underestimated. Table 3.1 highlights the ways in which social interaction dynamics can be taken into account when designing integrity policies.

Table 3.1. Precepts for strengthening integrity in social interactions in integrity systems

	Precepts for strengthening integrity in social interactions in integrity systems	See section
a)	Provide purposeful transparency	3.1
b)	Strengthen indirect reciprocity	3.1
c)	Strengthen integrity norms and identities	3.2
d)	Investigate and enforce strictly without creating a culture of fear and distrust	3.3
e)	Separate motivating and enforcing functions	
f)	Identify and counteract systemic de-stigmatisation of corruption (collective action trap)	3.4

A key factor in social dynamics is responsibility. When the burden of responsibility is spread over many individuals, the resulting diffusion of ethical accountability can pose an integrity risk. When several people jointly engage in misconduct, each of them feels less responsible, because guilt decreases when shared. These dynamics can undermine policies that aim to increase integrity by involving additional actors (e.g. open data

policies or the four-eyes principle). As such, it is important to avoid the diffusion of responsibility within an integrity system.

Additionally, transparency only serves integrity when those with responsibility have the capacity to act upon it. Efforts to share decisions with additional actors in a transparent manner need to be targeted and undertaken in a purposeful way. This is applicable to transparency efforts at all levels, whether they relate to opening up public data to millions of citizens or sharing a conflict-of-interest situation with a single colleague. If the relevant information is accessible, but the addressee is not aware or cannot act upon it, transparency does not ensure accountability.

Upholding the public interest can be understood as an act of indirect reciprocity, undertaken on the basis that others will do the same and the belief that integrity benefits all. Integrity policies can strengthen such behaviour by creating an institutional setting in which adherence to integrity is expected to be the norm. Setting a high standard of conduct is thus a key objective of integrity policies.

Enforcing this norm, in turn, is a balancing act between building trust in shared values and strictly investigating and sanctioning violations. Both functions gain credibility when they are separated (e.g. executed by different offices).

A group that has collectively diverged from a norm that supports integrity and has de-stigmatised a behaviour that enables corruption might find itself caught in a collective action trap. In this situation, an individual's motivation to act with integrity is no longer sufficient to break the vicious cycle and moral appeals will fall on deaf ears. Such situations need to be identified and systematically counteracted within an integrity system.

Chapter 4. Applying behavioural insights to integrity policies

Research on human behaviour and moral decisions provides an inspiring perspective on integrity policies. Chapter 4. invites policy makers to explore how behavioural insights can help shape the design of modern integrity policies.

Essentially, there are two approaches to applying behavioural insights to integrity policies:

1. Rethinking existing integrity policies and the design of integrity systems in the light of behavioural evidence. The behavioural approach addresses some of the discontent and criticisms surrounding traditional anti-corruption policies, such as the negative side-effects of overly strong control and enforcement approaches, while making an argument for promoting trust and an organisational culture of integrity. The lessons presented in the previous chapters have implications for the design of integrity systems and related policies. Section 4.1 thus maps out the relevance of behavioural precepts for integrity policies in line with the OECD Recommendation on Public Integrity (OECD, 2017[5]).

2. Designing a specific intervention to fix a behavioural pitfall, improve an existing policy or activate a certain behaviour. The application of behavioural insights through innovation interventions is gaining relevance for public policy making, as featured in the policy streams of the OECD, as well as those of the European Union and the World Bank. A 2017 OECD survey collected more than 100 examples of behavioural interventions in public policies (OECD[3]). Today, 135 public institutions worldwide have institutionalised behavioural capacities. The design of behavioural interventions is based on knowledge of human behavioural biases, cognitive limitations and social preferences. Often the application takes the form of a "nudge", a subtle change in the way a decision presents itself that helps individuals make a better choice to their own benefit. Section 4.2 examines how integrity policies can benefit from this trend and explores of the established behavioural interventions might be applicable to integrity policy making.

4.1. A human-centred perspective on integrity systems

The OECD Recommendation on Public Integrity provides a shared vision of the core elements of an integrity system. It shifts the focus from ad hoc integrity policies towards a comprehensive, risk-based approach with an emphasis on cultivating a culture of integrity across the whole of society (Figure 4.1). Examining integrity policy making through a behavioural lens places a different perspective on many elements of integrity systems.

Figure 4.1. The three pillars of the OECD Recommendation of the Council on Public Integrity: System, culture, accountability

Source: (OECD, 2017[5])

4.1.1. Commitment: Demonstrate commitment at the highest political and management levels within the public sector to enhance integrity and reduce corruption

The *tone at the top* matters for integrity. An unambiguous commitment to integrity sets the scene for how integrity is debated and perceived across the public sector and society. Defining, supporting, controlling and enforcing public integrity sends a clear message that integrity is valued and protected. In terms of the logic of indirect reciprocity, this can function as a motivation for others in the integrity system to uphold those values. Meanwhile, a clear and repeated commitment to integrity can serve to emphasise and raise individuals' moral reference points: it signals to all public officials that integrity is a crucial part of their professional identity.

Behavioural insights highlight the relevance of attitudes and perceptions for integrity. Integrity policy makers can shape these attitudes and perceptions by participating actively in the public debate. Timely and authentic communication can be a powerful complement to the policies of an integrity system. When a corruption scandal occurs, a public reaction from anti-corruption authorities can prevent discouragement and strengthen trust in the enforcement of integrity norms. Statements from political and institutional leadership create visibility for the actions undertaken in defence of integrity, while reassuring the public that the breach of an integrity norm is not silently accepted.

4.1.2. Responsibilities: Clarify institutional responsibilities across the public sector to strengthen the effectiveness of the integrity system

Shared responsibility can translate into shared guilt in cases of misconduct. As argued above, the diffusion of responsibility poses an integrity risk. Accordingly, clarification of responsibilities across the public sector not only increases the effectiveness of an integrity system, but can also strengthen the integrity of individual decision makers. The institutional design of an integrity system should ensure that responsibility for any type of choice and the related integrity risk is actively taken and perceived to its full extent. To this end, institutional and personal responsibilities need to be clearly defined.

Policies that add another actor to a process without giving this actor a unique responsibility, such as the four-eyes principle, do not function as effective tools to control corruption. The same can be said for policies for procedural clearance. A *de jure* requirement does not guarantee *de facto* implementation. On the contrary, the existence of a *de jure* requirement can relieve feelings of responsibility among those providing it as well as among those demanding it (Lambsdorff, 2016[75]). Demonstrating 'adequate procedures' or sufficient efforts for integrity cannot replace acting with integrity. Procedural clearance, formal requirements and reviews can never be allowed to replace responsibility for *de facto* integrity. Responsibilities within an integrity system should thus be not only clearly, but meaningfully, assigned.

4.1.3. Strategy: Develop a strategic approach for the public sector that is based on evidence and aimed at mitigating integrity risks

The OECD Recommendation on Public Integrity recognises the relevance of a strategic and evidence-based approach within the public sector. The growing empirical evidence gathered by behavioural science researchers is already feeding into strategic approaches based on sound evidence. From a behavioural perspective, the fact that time pressures and unfavourable working conditions limit people's capacity for self-control when confronted by opportunities for corruption (see section 2.2), argue for greater strategic planning and prioritisation within the public sector. Staff that are well-equipped, not overwhelmed and strategically guided are more resistant to the temptations of corruption.

However, empirical evidence shows that goals can also corrupt (Ordnez and Welsh, 2015[76]); (Welsh and Ordóñez, 2014[77]). Unrealistic goals create frustration and feelings of unfairness, which can serve as justifications for misconduct (see section 2.4). Faced with overly burdensome compliance rules, staff might cut corners to achieve performance goals and begin to tolerate non-compliance. Moreover, where performance evaluation is closely linked to the achievement of specific performance goals, staff might redirect all their efforts to achieving these performance goals rather than fulfilling their overall role. This often happens at the expense of overarching objectives – such as integrity.

A values-based strategic approach for the public sector –reflects integrity in the overarching strategy, as well as in management goals and performance appraisals. While the integrity-enhancing aspects of strategic management and performance evaluation are recognised, the potentially integrity-damaging aspects of over-compliance and frustration need to be acknowledged.

4.1.4. Standards: Set high standards of conduct for public officials

As can be concluded from Chapter 2. , standards of conduct are a behavioural reference: they signal a high moral reference point and contribute to developing professional identities for public employees. To help achieve the behaviour-changing potential of standards of conduct, the following points could be considered.

Standards and guidance for ethical conduct are often derived from a commitment to overarching values. Such values are the frame against which everyday choices and actions can be evaluated. As the number of items humans can store in their working memory is limited, a memorable set of values or key principles ideally has no more than seven elements (Miller, 1955[78]). Box 4.1 gives the example of the Danish Code of Conduct, which defines seven central duties to guide civil servants.

Box 4.1. Setting standards for integrity

Meaningful and memorable public sector codes

The Australian Public Service (APS) Values

In 2010, the Advisory Group on Reform of the Australian Government Administration released its report, which recognised the importance of a robust values framework to a high-performing, adaptive public service, and the importance of strategic, values-based leadership in driving performance. The APS values aim to provide "a small[er] set of core values that are meaningful, memorable, and effective in driving change". The model follows the acronym "I CARE":

- **I**mpartial
- **C**ommitted to service
- **A**ccountable
- **R**espectful
- **E**thical

The Colombian Integrity Code

In 2016, the Colombian Ministry of Public Administration initiated a process to define a General Integrity Code. Through a participatory exercise involving more than 25.000 public servants through different mechanisms, five core values were selected:

- Honesty
- Respect
- Commitment
- Diligence
- Justice

In addition, each public entity has the possibility to integrate up to two additional

values or principles to respond to organisational, regional and/or sectorial specificities.

Syv centrale pligter – Seven key duties

The Danish Agency for Modernisation (MODST) under the Ministry of Finance issued the "Kodex VII" – a Code of Conduct for Danish civil servants. The Code defines seven central duties: (i) Legality, (ii) Truthfulness, (iii) Professionalism, (iv) Development and co-operation, (v) Responsibility and management, (vi) Openness about errors and (vii) Party-political neutrality.

The "Kodex VII" describes the relevance and the implications of each duty for the Danish public sector. Moreover, MODST provides fictional case studies that can be used to practice the application of "Kodex VII". Potential solutions for the case studies are available to public institutions, but are not published.

Sources: (Ministry of Finance, 2015[79]), Australian Public Service Commission (2011), "Values, performance and conduct", http://www.apsc.gov.au/about-the-apsc/parliamentary/state-of-the-service/state-of-the-service-2010/chapter-3-values,-performance-and-conduct; Australian Public Service Commission (2012), "APS Values", www.apsc.gov.au/aps-employment-policy-andadvice/; http://www.funcionpublica.gov.co/eva/codigointegridad/index.html

In addition, standards of conduct are ideally defined through a participative process to ensure that the result is meaningful and actionable for public employees. The overarching values, the language used and the examples provided in a code of ethics need to be able to activate and remind the user of their own moral reference point. The impact on the user will be even greater if the values, language and examples used are personally meaningful and relatable.

The dialogue on standards of conduct does not end after the development of a code. Regularly refreshing one's commitment to ethical behaviour adds value to the dynamics of moral choice making. Within an integrity system, ethics training and regular dialogue in team meetings could serve this function. Integrity policy makers could also consider approaches inspired by behavioural insights to provoke ideas and discussions on integrity, such as e-mail reminders or competitive quizzes. Section 4.2 might provide some inspiration for such focused interventions.

Public employees who have familiarised themselves with the standards of conduct through training and discussions and have internalised the related values can be encouraged to mentally reactivate their commitment to these standards through regular reminders. References to core public sector values in the work place and in official communications can help to strengthen ethical reflection and serve to strengthen moral choices (see Chapter 2.). While a mousepad or a pen with the word *integrity* on it might not seem likely to have a strong impact on a person's behaviour, they can act as moral reminders and alter actions by establishing a link with previous discussions and a sense of personal commitment. The same can be said for the message "thank you for your honesty" or the honour code in the experiments described in Box 2.3 and Box 3.2.

Ethical reflections making a regular appearance in the everyday routines increase the chance that unethical behaviour is identified and spoken up against. Moral reminders are one way for policy makers to trigger ethical reflections. The research on reminders shows that they work best when introduced in a timely manner. Integrity policy makers could therefore aim to trigger ethical reflection during at-risk procedures. A signature, for example, can function as a tool to invoke stronger personal ethical engagement. Research

shows that a signature has a stronger impact on choices when it is required prior to the decision that poses an integrity risk (Shu et al., 2012[80]).

4.1.5. Whole of society: Promote a whole-of-society culture of integrity, in partnership with the private sector, civil society and citizens

Active dialogue with citizens and regular contact with those affected by public sector decisions amplifies feelings of ethical responsibility within the public sector. Breaking out of a limited sphere of interaction can also help public sector decision makers to feel that their work has implications for people's lives and reduces the perception that corruption is a victimless crime (see section 2.5). Furthermore, once public sector staff can put a face or a story to the people who place trust in them, they might take more pride in protecting this confidence. An opportunity to engage with stakeholders is often also an opportunity to increase integrity. Otherwise, public employees who only interact with those affected by their work through complaints may become frustrated and eventually lose their motivation to uphold the public interest.

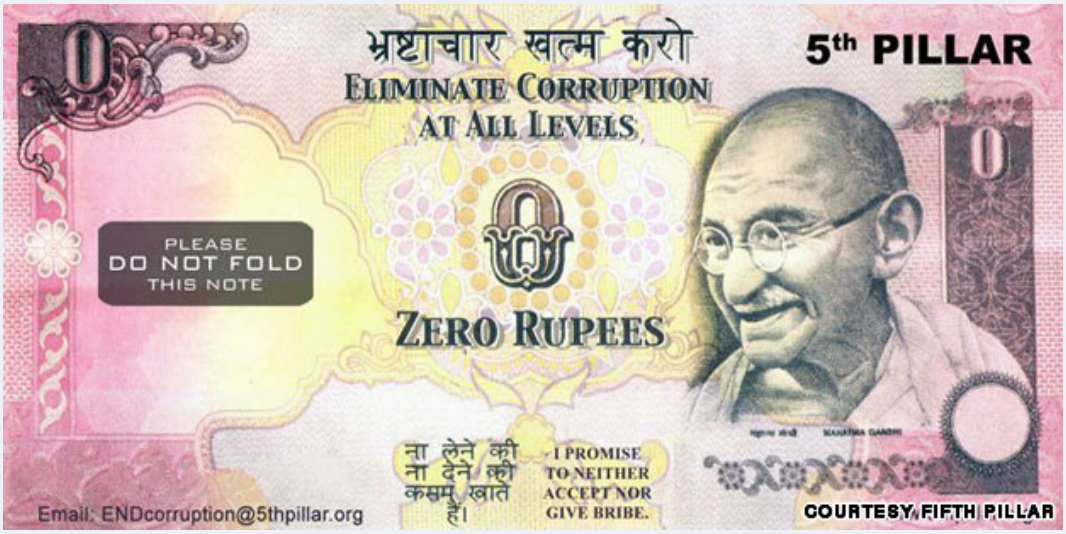
One specific behaviourally informed way to encourage feedback from citizens is the so-called ambient accountability mechanism. This takes the form of a physical element such as a poster or a screen, which is placed directly in the public office for citizens to leave comments (Zinnbauer, 2012[81]). Box 4.2 highlights a particularly innovative approach to obtaining direct citizen feedback, in this case regarding corrupt public officials. In addition to physical feedback, policy makers could embrace the digitalisation of services and procedures as an opportunity to integrate feedback questions.

Box 4.2. The zero rupee note

The non-governmental organisation (NGO) 5th Pillar created a zero rupee note in India with the inscription, "I promise to neither accept nor give a bribe". The organisation has distributed more than 3 million such notes among the public, mostly through students in schools and colleges, all of whom have taken a verbal pledge against corruption.

Handing over the note is a surprising and revealing way to refuse co-operation in a corrupt transaction. Dissatisfied recipients would have to break the silence by initiating a debate in which they do not have the moral high ground. The NGO reports that thousands of citizens who have handed out the notes found that their counterparts immediately yielded their request without insisting on the bribe. One official was supposedly "so stunned to receive the note that he handed back all the bribes he had solicited for providing electricity to a village" (World Bank, 2015[4]).

Figure 4.2. Zero Rupee Note



Sources: 5th Pillar - Citizens for Democracy (n.d.), Zero Rupee Note (ZRN), https://5thpillar.org/programs/ (accessed on 23 January 2018). (Banerji, 2016[82]); (World Bank, 2015[4]).

4.1.6. Leadership: Invest in integrity leadership to demonstrate public sector commitment to integrity

As mentioned above, the tone at the top matters for integrity. However, the top does not refer only to the highest political and management levels. The relevance of middle and lower management cannot be downplayed as their immediate impact on the staff under their direct responsibility may arguably be even higher and more direct. They can help to raise the moral reference point of their employees by promoting discussions about ethical dilemmas or by setting an example with respect to day-to-day conduct in accordance with the shared values of the public service.

Additionally, the preventive approach of building trust and encouraging integrity is ideally implemented by managers other than those responsible for the detection, enforcement and control of corruption (see section 3.4). Managers that have ideally

established a trusting relationship shaped by positive reciprocity are not always the most likely to detect improper conduct among staff. People tend to ignore information contradicting their existing beliefs, sometimes dismiss warning signals and are less likely to judge the behaviours of those close to them as unethical, behavioural research shows (Gino and Galinsky, 2012[83]); (Eil and Rao, 2011[84]); (Hildreth, Gino and Bazerman, 2016[85]). In addition, the enforcement integrity among staff often involves communicating distrust. Managers who implement strict measures of control, or more severe, sanction their staff, risk crowding out their intrinsic motivation and losing their trust (see section 2.7).

4.1.7. Merit based: Promote a merit-based, professional public sector dedicated to public service values and good governance

Integrity fundamentally concerns human behaviour. As such, the civil service and human resource management play a key role in its promotion. Per definition, integrity is indirectly part of every job description in the public sector. Accordingly, "I'm only doing my job" is never an excuse for unethical behaviour on the part of any public official. Behavioural research has long understood that people tend to act according to the roles they are given. Priming a certain social identity can lead people to act according to the characteristics of this identity (Benjamin, Choi and Strickland, 2010[86]).

Public employers need therefore to clearly communicate to their staff that they are expected to uphold the public interest and give them actionable guidance on how to do so, for instance through regular integrity training. Where integrity is clearly linked to the professional identity associated with a position in the public sector, acting with integrity becomes an act of pride and duty. The prominence of integrity for the role can be signalled to new applicants for public sector positions already in job postings and form an integral part of the selection and recruitment process. The Office of Government Ethics in the United States requires all written offers of employment in the executive to include a statement regarding the agency's commitment to ethics and to set role-specific expectations early. New supervisors are alerted of their special responsibility for advancing government ethics. Box 4.3 provides an example from Australia on how to integrate integrity into selection processes.

Box 4.3. Recruitment processes and integrity – experience from Australia

"Filters" can be built into a recruitment process to ensure applicants are suited to an organisation's requirements. In Australia, one agency analysed disciplinary issues among new recruits after 12 months on the job and identified a need to better manage integrity indicators earlier in the selection process. Interventions were then instituted at key stages:

- A question and answer survey was incorporated into the general information for potential applicants. It asked questions about how candidates felt about certain working conditions and interactions. Based on an indicative score, potential applicants were encouraged to proceed to the next stage or to speak about the role with people who knew them well before proceeding to the next stage. This approach encouraged self-filtering by applicants.
- The application asked more targeted integrity questions about the background and experiences of applicants (e.g. dealing with authority, diverse cultures and financial management). These provided base data for comparative purposes.
- Successful applicants in the technical assessment phase were asked to retake the integrity questions. Experts were then asked to identify discrepancies or anomalies between the datasets and to follow these up individually with applicants. The delay between administering the questions increased the validity of the data.
- Only applicants who passed both the technical and the integrity phases were invited for face-to-face interviews, which included a practical role play.

The outcome was a considerable decrease in disciplinary issues and increased retention rates for new recruits.

Source: Australian Merit Commission, June 2016.

However, the understanding that public employees have with respect to their own role and professional identity can diverge widely from that of the management. Conducting regular public sector staff surveys and encouraging ethical reflections can contribute to maintaining awareness of integrity as part of the identity of public employees.

The relationship between a public employer and its employees is more than an exchange of labour for payment; it is shaped by reciprocity. Employees who receive a generous wage are willing to work harder, even where this is not observable (Akerlof and Yellen, 1990[87]); (Gilchrist, Luca and Malhotra, forthcoming[88]); (Van Veldhuizen, 2013[89]). Their behaviour is shaped by the way they are treated. Non-monetary incentives are even more effective in activating intrinsic motivation (Gneezy, Meier and Rey-Biel, 2011[90]); (Kosfeld, Neckermann and Yang, 2014[91]); (Kosfeld and Neckermann, 2011[92]). Employees who feel they are trusted, appreciated and recognised will be motivated to reciprocate through strong performance (Falk and Kosfeld, 2006[44]); (Fehr and List, 2004[93]). Employees who are treated with integrity are more likely to uphold this value in their daily work.

Individuals who are failed by an institution, a norm or their superior might eventually become frustrated and stop contributing to the public good of integrity. Employees who feel treated unfairly could feel licensed to act unethically (see Chapter 2.). Feeling

mistreated not only reduces the obligation of reciprocity towards the employer and the employer's norms, but acts as a gateway to justifications for misconduct (see section 2.4). For example, an individual disadvantaged during a hiring process because another applicant exploited personal connections, might lose trust the integrity of the system and decide to rely on connections themselves in the future. If the number of people who have lost trust rises, the integrity norm will start to fall apart. Organisational fairness is a crucial condition for integrity (OECD, 2009[94]).

Modern managers therefore seek to establish a trusting and appreciative relationship with their employees. Integrity should be a visible part of this relationship. Internal communication, goal setting and role descriptions could include integrity as a core objective and subject of the employer-employee relationship. Consequent communication of ethical standards makes clear that integrity is expected of the members of the organisation, who can in turn expect the organisation to respect integrity as an organisational value. Evidently, such claims must be applied in practice. Employees can be discouraged to uphold norms and principles that they feel are not respected throughout the organisation (see Chapter 3.).

In an open organisational culture that aims to prevent integrity breaches, it is thus crucial that staff feel they are treated fairly, that errors are tolerated and that they have the opportunity to voice their concerns should that change. Ensuring that integrity is appreciated and that speaking up is not penalised promotes integrity as a reciprocating norm. Settings in which those who adhere to the rules have to bear additional costs (e.g. increased paperwork, conflict with colleagues, longer waiting times) should be avoided. As noted above, organisations could monitor the perceptions of staff through regular employee surveys. Employee complaint mechanisms could also be in place to ensure conflicts are addressed. Any suppressed feelings of sacrifice or under-appreciation among employees will pose an integrity risk.

4.1.8. Capacity building: Provide sufficient information, training, guidance and timely advice for public officials to apply public integrity standards in the workplace

Integrity training offers a unique opportunity to emphasise and reinforce shared values, and can thus emphasise and contribute to raising the moral reference point of participants. Induction trainings, for instance, which take place at the very start of a public servant's career, communicate the values of the respective public organisation and provide guidance on the expected behaviour in the social context. Training well aligned with the day-to-day reality of the organisation face a lower risk to be perceived as naïve or to be seen merely as a formal requirement. For sustainable capacity building, training needs to be repeated over time, and may be adjusted to the needs of specific target groups such as at-risk positions or management.

Interactive components where participants are confronted with realistic situations are more likely to generate a personal mental commitment to integrity than mere presentations by trainers (Bazerman and Tenbrunsel, 2011[20]). Box 4.4 provides the example of dilemma training implemented by the Flemish Government in Belgium. The design of training can also build on behavioural insights. For instance, training can benefit from confronting participants with past situations where they likely behaved in an unethical way. The participants then need to resolve this moral dissonance by performing a positive action.

Box 4.4. Dilemma training in the Flemish Government (Belgium)

The Agency for Government Employees in the Flemish Government offers dilemma training to public officials. The training provides the participants with practical situations in which they face an ethical choice with no clear path to resolving the situation with integrity. The facilitator encourages discussion between the participants about how to resolve the situation and helps them explore the different choices. The focus of the training is the debate rather than possible solutions, as the objective is to help participants identify how different values might act in opposition to one other.

In the majority of training sessions the facilitator uses a card system. The participants receive four "option cards" with the number 1, 2, 3 or 4. A set of "dilemma cards" are then placed on the table. Each dilemma card describes a situation and gives four options for resolving the dilemma. In each round, one of the participants reads out the dilemma card and the options. Each participant indicates their choices using their option cards and explains the motivation behind their choice. The participants then discuss the different choices. The facilitator remains neutral, encourages the debate and suggests alternative ways to look at the dilemma (e.g. sequence of events, boundaries for unacceptable behaviour).

One example of a dilemma situation could be:

"I am a policy officer. The Minister needs a briefing within the next hour. I have been working on this matter for the last two weeks and should have already been finished. However, the information is not complete. I am still waiting for a contribution from another department to verify the data. My boss asks me to submit the briefing urgently as the Chief of Cabinet has already called. What should I do?

1. *I send the briefing and do not mention the missing information.*
2. *I send the briefing, but mention that no decisions should be made based on it.*
3. *I do not send the briefing. If anyone asks about it, I will blame the other department.*
4. *I do not send the information and come up with a pretext, and promise to send the briefing tomorrow.*"

Other dilemma situations could cover themes such as conflicts of interest, ethics, loyalty, leadership and so on. The training sessions and situations used can be targeted to specific groups or entities.

Source: https://overheid.vlaanderen.be/omgaan-met-integriteitsdilemmas (in Dutch).

Beyond training, the OECD Recommendation on Public Integrity also emphasises the importance of the availability of guidance when public officials are faced with specific doubts or ethical dilemmas. Although integrity is ultimately the responsibility of all organisational members, dedicated "integrity actors" can play a particularly important role in stimulating integrity and shaping ethical behaviour (OECD, 2009[94]). In addition, guidance on ethics and conflict of interest in cases of doubts and dilemmas can respond on an ad hoc basis when public officials are confronted with a specific problem or doubts. Written guidance, such as the REFLECT model provided in Australia (Box 4.5), together with complementary trainings can support public officials in making the correct moral choices.

Box 4.5. Guiding public officials facing ethical dilemmas in Australia

To support the implementation of ethics and its integrity regime, the Australian Public Service Commission has enhanced its guidance on the Australian Public Service (APS) Values and the Code of Conduct. This includes integrating ethics training into learning and development activities at all levels.

To help public servants address ethical dilemmas during the decision-making process, the Australian Public Service Commission developed a decision-making model, known as REFLECT:

1. **RE**cognise a potential issue or problem. Public officials should ask themselves:
   - Do I have a gut feeling that something is not right or that this is a risky situation?
   - Is this a right vs right or a right vs wrong issue?
   - Recognise the situation as one that involves tensions between APS Values or the APS and their personal values.
2. **F**ind relevant information
   - What was the trigger and circumstances?
   - Identify the relevant legislation, guidance, policies (APS-wide and agency-specific).
   - Identify the rights and responsibilities of relevant stakeholders.
   - Identify any precedent decisions.
3. **L**inger at the "fork in the road"
   - Talk it through, use intuition (emotional intelligence and rational processes), analysis, listen and reflect.
4. **E**valuate the options
   - Discard unrealistic options.
   - Apply the accountability test – public scrutiny and independent review.
   - Be able to explain your reasons/decision.
5. **C**ome to a decision
   - Come to a decision, act on it and make a record if necessary
6. **T**ake time to reflect
   - How did it turn out for all concerned?
   - Learn from your decision.
   - If you had to do it all over again, would you do it differently?

Source: (Godwin, 2009[95])

4.1.9. Openness: Build an open organisational culture within the public sector responsive to integrity concerns

Integrity breaches often remain unchallenged in their immediate environment due to the strong behavioural forces of reciprocity, the power of unconscious biases, and the

tendency towards conformity of judgement and behaviour within groups. Corruption is often the end of a slippery slope consisting of small misjudgements and gradual acceptance. Moreover, research points to the fact that leaders and those who identify strongly with their organisation present an even higher risk of letting integrity breaches go unchallenged. (Barkan, Ayal and Ariely, 2015[15]); (Kennedy and Anderson, 2017[96]). Managerial oversight is therefore insufficient as the only defence of integrity.

The ability to voice concerns in the face of misconduct is key to upholding an integrity system. Public employees who witness wrongdoing are expected or required to report it. However, employees may be tempted to turn a blind eye to irregularities to avoid trouble for their colleagues or organisation. Behavioural insights can help to understand and overcome this dilemma.

Speaking up requires a strong organisational culture. In organisations lacking a culture of openness to discuss ethical concerns or dilemmas, whistleblowers put their personal relationship with their colleagues and even their job at risk to protect the public interest. A team member who voices concern about the behaviour of another team member questions this colleague's personal integrity. Regardless of whether the disclosure is based in fact, it may affect mutual trust among team members (Reuben and Stephenson, 2013[97]). Whistleblowers are sometimes also suspected of personally benefiting from the situation, for example, by wanting to inflict harm on the person they are reporting. In cases where direct and indirect reprisals are not prevented or mitigated, employees will need courage and a strong moral compass to speak up in the light of misconduct.

However, protecting the public interest may require speaking up against a respected colleague, an established practice or an accepted system. Public sector employers emphasise integrity hoping that their staff will voice their concerns, but it would be naïve to rely solely on the categorical morality of public officials under such circumstances. As noted in Chapter 2. , moral choice is a dynamic process in which people also consider their own self-interest. Whistleblower policies stand to gain by acknowledging this and aim to reduce the burden a whistleblower has to carry to the extent possible.

From a behavioural perspective, the following two aspects can facilitate speaking up:

- **Certainty**: Potential whistleblowers will evaluate the act in light of what follows. Uncertainty about the procedure or the consequences of reporting makes it harder to take this step. Whistleblowing policies should thus clearly and credibly state the process of reporting, make this information known to all staff, and make those who are responsible for handling disclosure of misconduct accountable for their actions. In particular, whistleblowers should be provided with as much certainty as possible regarding whether they will be granted protection and whether their case will be investigated. Furthermore, whistleblowers will speak up with more confidence if they know which behaviours constitute misconduct. A broad definition of the types of misconduct eligible for reporting and knowledge and awareness of this definition could increase use of the whistleblowing system.
- **Confidentiality**: Just as control can act as a burden on the trust between managers and their team, voicing an integrity-based concern can burden a public employee's relationship with their managers and colleagues. Public officials might choose to avoid risking longstanding work relationships, especially on the basis of suspicion. Protecting the identity of whistleblowers, for example, through a third-party reporting mechanism or confidential digital reporting system, allows them to seek advice on an issue without disrupting their team relationships. Once they have

gained confidence that the incident in question is indeed unethical, and have greater certainty regarding the consequences, they might choose to reveal their identity.

Once again, a dedicated integrity contact point, either a unit or a person, can be conducive both to certainty and confidentiality. An external person serving as contact point could provide ethical guidance and procedural advice to public officials, while keeping their identity confidential. With this support, the public official can make an informed decision about disclosing misconduct. When an external contact points is well-known, trustworthy and clearly separated from enforcement functions, this increase the likelihood that people will turn to it.

Protection from reprisal or effective remedies, interim relief, legal support or honorary awards can be other ways to show appreciation of a whistleblowers dedication to integrity. Moreover, thorough investigation of the reported case and sanctioning of misconduct honour integrity. If there is no certainty that the reported cases will be consequently followed-up with, then taking on the cost of reporting is not reasonable for a potential whistleblower.

The logic of reciprocity (see Chapter 3. calls for whistleblowing to be rewarded. By exposing behaviour that can potentially harm the organisation as well as the public interest, whistleblowers mitigate financial and reputational risks within an organisation. However in an organisation where the most members are generally committed to integrity appreciation for whistleblowers does not necessarily need to be expressed through a monetary reward. In fact, a monetary reward can in some contexts even discourage speaking up. Incentives can have a hidden cost, that is, a counterproductive effect on a trust-based relationship (Fehr and List, 2004[93]). Monetary incentives might make the actions of the whistleblower appear to be based on self-interest rather than principle. People's estimation of a moral choice is generally increased when the individual in question acts without direct reward (Hoffman, Yoeli and Nowak, 2015[98]) see also section 3.1). Nonetheless, an organisation within a collective action trap (see 3.5) where norms are already deterred and whistleblowers fear severe vengeance can use monetary rewards as an incentive to break the silence in corrupt networks.

As Chapter 3. emphasised, integrity policy makers could seek to strengthen indirect reciprocity. In this sense, an integrity system encourages whistleblowers to act with integrity not for personal monetary reward (direct reciprocity), but to protect the shared norm. Nonetheless, a person who places much at stake to protect integrity will usually hope that others value this dedication to the public interest (indirect reciprocity). Such appreciation can be integrated in whistleblowing systems (e.g. by awarding whistleblowers) and wider human resource management policies (e.g. through positive recognition of whistleblowing in performance appraisal).

Ways of showing appreciation of a whistleblower's dedication to integrity include protection from reprisals or effective remedies, interim relief, legal support and honorary awards. Integrity is also honoured through thorough investigation of the reported case and punishment of misconduct. If there is no certainty that the reported case will be investigated, the cost of reporting might not be perceived as reasonable by a potential whistleblower.

Sometimes employees do not report misconduct simply because they do not perceive it as such. Conformity bias (Asch, 1952[63]); (Asch, 1951[99]), self-serving social identity and justifications can prevent people from becoming aware of ethical misconduct within their own work environment (see sections 2.4 and 0). A behaviour that is accepted by

colleagues or managers is less likely to be identified as unethical. Integrating regular ethical reflection into public officials' work routines can counteract such effects. This could take the form of a moral reminder message, a refresher ethics workshop or a regular discussion of integrity topics with an external third party. Most importantly, these activities need to create salience for ethical standards and invite public officials to reflect on whether these standards are being upheld in their own work.

Finally, whistleblowers who are themselves involved in the misconduct in question might be deterred by the prospect of sanctions. Will the team's reputation be destroyed or all past work declared void? Will they personally be under scrutiny? Box 3.3 in Chapter 3. describes a case where an actor was scared to report misconduct because he had unintentionally become involved. Fear of sanctions can create confidence among corrupt partners that no-one involved will sever the reciprocal arrangement. Asymmetric liability in whistleblower law can break this dynamic of extortion. Mitigation or limited sanctions for confidents who speak up can encourage such insiders to become whistleblowers (Abbink et al., 2014[100]); (Apesteguia, Dufwenberg and Selten, 2007[101]); (Buccirossi and Spagnolo, 2006[72]); (Lambsdorff and Nell, 2007[73]).

4.1.10. Risk management: Apply a control and risk management framework to safeguard integrity in public sector organisations

Supporting ethical conduct without surrendering control safeguards is a central challenge for any integrity system. Internal controls, audits and risk assessment policies are key policies concerned with monitoring and managing integrity risks. Sometimes these controls can be automatised and thereby delinked from human behaviour. In many cases, however, those controlling and inducing the risk are humans. Behaviourally informed control and risk management policies designed to anticipate human ethical blind spots can help reduce errors.

Human error can be a challenge for assessments that rely solely on human judgement. However, humans often err systematically; thus, certain misjudgements can be anticipated and avoided. For example, it has been established that humans struggle to correctly assess probabilities (Kahneman and Tversky, 1979[102]). Estimating both the likelihood and consequences of an event at the same time is cognitively challenging. Faced with such complex estimations, people tend to let one or the other guide their judgement. Focusing on the probability, an assessor might underestimate an integrity risk, because they think the event is unlikely to occur. When judging the probability of an incident, risk assessments could be deterred by the so-called representative heuristic (Kahneman and Tversky, 1979[102]): people tend to associate higher probabilities to events they can imagine or have experienced. A person might assume an integrity risk is very unlikely to occur, simply because they have never personally experienced it or because it has apparently never happened in the organisation before.

In response to this, the subject matter of any assessment could be broken down and stripped of all redundant information that could bias the decision maker. A risk assessor could, for example, estimate the impact of an event independent from its probability of occurring. Furthermore, auditors and other individuals that regularly conduct independent assessment could be made aware of potential bias and be trained to counteract it.

Instead of asking people directly about their own integrity risk, it might be better to pose general questions or ask questions about a similar person in a similar organisation. When people are asked to assess their own integrity risks, they are bound to be biased in favour of themselves (see Chapter 2.). Positive moral self-assessment clouds their judgement,

leading them to underestimate integrity risks related to their own role, as compared with the judgement of an outside observer. Along the same line, leaders might overestimate their own ability to prevent an integrity breach. People with high trust in their organisation might underestimate the likelihood of an integrity breach among their colleagues. Posing self-assessment questions more indirectly can reduce overconfidence and social desirability bias. For example, an interviewer could extract sensitive information by asking, "How likely do you think it would be for someone in a similar position to have done this?"

Certain roles have higher integrity risks and are therefore more closely examined in risk assessments and internal audits. Behavioural research can provide a few insights to inform the criteria for identification of these roles:

- People with power tend to judge their own integrity breaches more leniently (Barkan, Ayal and Ariely, 2015[15]).
- Higher ranks statistically identify more with their organisation. This weakens their ability to identify an integrity breach among their own colleagues (Kennedy and Anderson, 2017[96]).
- People who are at risk of exclusion from groups are more likely to act in an unethical manner for the sake of the group (Derfler-Rozin, Pillutla and Thau, 2010[103]); (Thau et al., 2015[104]).
- As discussed in section 2.5, the perceived immorality of an action decreases with distance from the potential victims of an action. Individuals whose corrupt choices impact people they are never likely to meet thus have a higher integrity risk than those who interact closely with those whose lives they impact regularly.

4.1.11. Enforcement: Ensure that enforcement mechanisms provide appropriate responses for all suspected violations of public integrity standards by public officials and all others involved in the violations

Detection and sanctioning of corruption help build trust (see section 3.4). Enforcement of formal norms gives those committed to integrity a sense of fairness. Separating such enforcement from encouragement allows both functions to achieve their full psychological strength (Lambsdorff, 2012[55]): Human resource policies that are by nature distrustful (e.g. reporting requirements, auditing, checking attendance) are more convincing when communicated separately from policies built on a trusting relationship (e.g. performance measurement, code of ethics). Ideally, such policies are introduced by different institutions or individuals with different roles. Integrity as a subject of dialogue between manager and staff can then be framed in a positive, motivating manner. For example, integrity can be communicated as part of organisational identity and as a performance goal.

In turn, it is important to prevent a perception of impunity as this would undermine the credibility of rules and procedures, promote cynicism among staff and provide individuals with arguments to justify their own integrity breaches (see section 2.4). As emphasised in section 3.5, organisations entrapped in a collective action trap, where it is rational for individuals to adapt to the corrupted context, require strict enforcement in order to change the organisational culture. Thorough investigations of integrity breaches, credible sanctions and impartial enforcement signal a valuation of integrity and provide certainty that laws and regulations cannot be circumvented without punishment (see sections 3.4 and 3.5). Indeed, it is important to emphasise that sanctions are not necessarily limited to criminal sanctions. According to the severity of the infringement, a

wide spectrum of possibilities is available to provide negative feedback, ranging from informal social sanctions, to disciplinary and criminal sanctions.

4.1.12. Oversight: Reinforce the role of external oversight and control within the public integrity system

External auditors examine facts and numbers in accordance with clear-cut guidelines and standards. They are also subject to strict controls to ensure they do not falsify their work. However, sometimes "good accountants do bad audits" (Bazerman, Moore and Loewenstein, 2002[105]): Despite having the external appearance of a hard science, auditing leaves considerable leeway for leniency and ambiguity. Wherever facts depend on interpretation there exists the possibility for unconscious bias.

Auditors can be made aware of and trained to counteract biased judgement. Sections 2.3 and 2.4 describe how mental shortcuts and misperceptions can lead humans to make biased judgements and yield to unethical behaviour without even being aware of their actions. This phenomenon is often referred to as "ethical blindness" (Bazerman and Tenbrunsel, 2011[20]); (Sezer, Gino and Bazerman, 2015[106]). Auditors and other individuals responsible for oversight functions could make use of behavioural insights and evidence to compensate for such unintentional blind spots. Although knowledge about behavioural biases alone is not sufficient to ensure they will have no impact on others, an awareness of ethical blindness allows decision makers to strengthen their resilience against typical sources of bias, such as overconfidence, conformity, justifications and representativeness bias.

Individuals performing auditing roles can protect their impartiality by performing only auditing functions. Providing consulting or standard-setting services in addition to auditing can challenge the objectivity of auditors, an issue implicitly recognised by the International Organization of Supreme Audit Institutions (International Organization of Supreme Audit Institutions, 2010[107]). In one role they flag problems, while in the other they design solutions. To avoid conflict of interest, auditors could refrain from performing both tasks for the same organisation or assessing a solution they designed or implemented themselves. Even where conflict of interest is well-managed, the psychological challenge of a double role can remain and potentially result in unconsciously biased judgements.

Consequently, clearly separating internal control activities from internal and external audit functions could strengthen their objectivity. While the responsibility for management of integrity risks and functions of internal control is normally shared within an organisation, internal audits are clearly a checking and monitoring function. Autonomy and impartiality – also mentally – are essential to these functions. In addition, perceived close links between audits and the internal control environment could harm the sense of ownership of internal control and risk management among public managers.

4.1.13. Participation: Encourage transparency and stakeholder engagement at all stages of the political process and policy cycle to promote accountability and the public interest

Implicit biases and the intuitive tendency to reciprocate can create an unconscious bias (see section 2.2). Judgments considered to be objective by those who make them might actually distorted by misperceptions. The risk of such unconscious bias increases with psychological closeness, resulting in a tendency to favour people similar to oneself or

with whom one interacts frequently. Where unconscious bias functions as an entry point for undue influence, it becomes an integrity risk. This is particularly relevant for individuals expected to make impartial decisions, as is often the case for policy makers and regulators.

Frequent interaction with different stakeholders is often essential for the work of key decision makers. Regulators, for example, interact with ministries, parliaments, regulated industries and citizens, who are the ultimate beneficiaries of governmental and regulatory actions. "These interactions are inevitable and desirable" (OECD, 2016[108]), but they can also impact decision making. People use mental shortcuts, such as the "availability heuristic" (Tversky and Kahneman, 1973[109]), according higher relevance to things that they can recall, put an example to or have recently experienced. Thus, without any exertion of undue influence, a frequent exchange with a particular group can bias decisions towards this group's interests.

A balanced consideration of interest cannot be achieved by prohibiting such exchanges, but rather by ensuring the equal representation of interests in the minds and perceptions of decision makers. It is therefore important to ensure a level playing field in terms of access to the policy-making process and the policy cycle for all interest groups (OECD, 2017[31]).

4.2. "Nudging" for integrity

Behavioural insights and their use in policies have been popularised through low-cost, small interventions, often called "nudges". Such interventions are now being applied in public policies around the world (OECD, 2017[3]). Their use is usually intended to overcome a particular behaviour-related challenge within a policy design (e.g. when people forget to sign up for a programme).

However, integrity is too complex and broad an issue in behavioural terms to be defined as the intended outcome of a "nudge". Nonetheless, behavioural interventions can be used to increase the effectiveness of integrity policies or procedures within an integrity system. The first step is to identify challenges related to behavioural concepts within an integrity system. Behavioural interventions can then intervene wherever the success of a policy or procedure is compromised by subjective perceptions and judgements on the part of decision makers.

The way in which choices present themselves can alter the outcomes of a decision. For example, humans are generally averse to losses. As a result, most people behave differently in the face of a potential loss compared to a potential gain. This preference also affects ethical choices: experiments showed that people were more likely to act unethically to prevent a loss than they would to make a gain (Kern and Chugh, 2009[110]). The amount that could be won or lost was the same in both cases – the difference was merely a matter of phrasing.

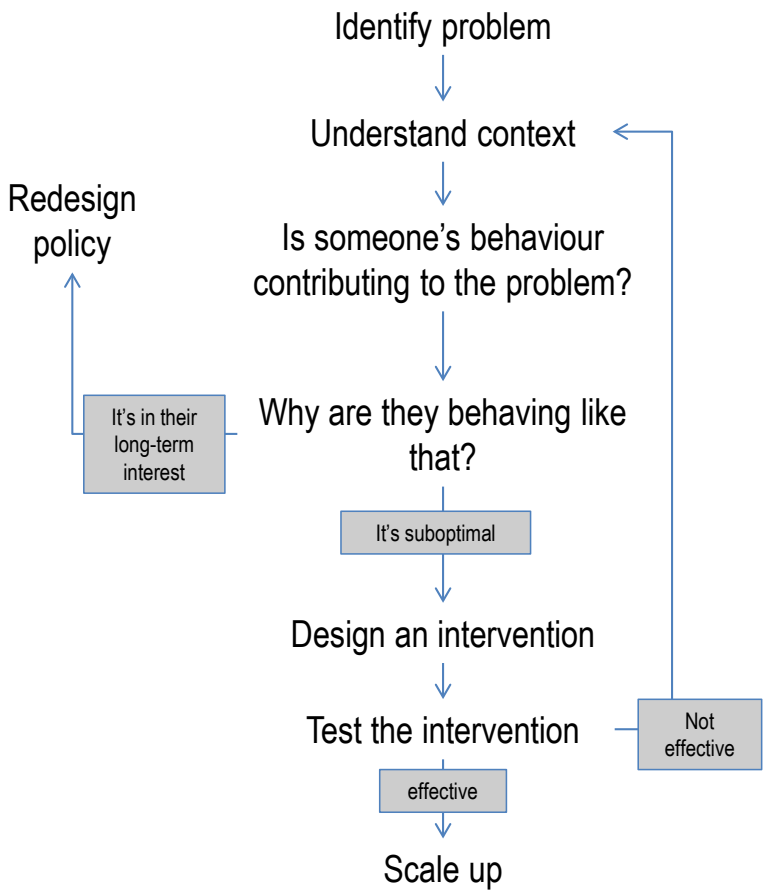
A nudge alters the choice architecture – the way in which people are presented with a choice – without limiting the options involved. It builds on behavioural research mapping the various biases in human perception and decision making. By considering all the ways in which human choices differ from rational predictions, the choice architecture can help people make a good choice in their own interest. Typical alterations to the choice architecture include (OECD, 2017[111]):

- changes to wording, naming and communication

- changes to the physical environment
- changes to procedures and forms
- use of feedback and interaction
- rewards, goal setting and commitment.

The identification of a potential area of application for behavioural interventions requires a detailed analysis of the problem to be solved. Figure 4.3 illustrates this process. Not every integrity challenge can be fixed with a behavioural intervention, but behaviours are involved in almost any integrity policy. Whenever this behaviour is "suboptimal" – meaning that the choices of the individuals might be falling short of achieving their own best interest – a behavioural intervention can help.

Figure 4.3. Targeting a behavioural intervention



For example, assume that an organisation has a confidential integrity contact point, but that this person or unit is not often contacted regarding advice about potential integrity breaches. If the underlying reason is that the contact point does not provide competent advice, it is in the interest of people not to approach them. Instead of a behavioural intervention, this situation requires a policy re-design to rethink or strengthen the capacities of the contact point. However, it is also possible that people may not be aware of the existence of the contact point. In such cases, a behavioural intervention could be used to remedy the situation by raising awareness of the existence of the contact point.

Where the concerned behaviour is central to the policy and a strong bias or blindness is affecting the choice, the impact of a small and cost-effective behavioural intervention can be significant. Although behavioural interventions are no panacea, it is worth investigating their potential applications. In cases where behavioural insights cannot necessarily be used to design a whole new policy, they can still make conventional policies more effective. For example, behavioural insights will probably not yield a whole new approach to asset declarations, but they could be used to nudge people to declare their assets on time or to make declaration forms more user-friendly.

Behavioural policy making has developed a considerable repertoire of policy responses to encourage certain choices and counteract biases. Common tools for choice architecture changes include the following:

- **Default settings**: People show a certain inertia with regard to making active choices. This is why one of the most powerful nudges is to change the default (e.g. default inclusion of travel insurance when buying flight tickets).
- **Reminders**: The human mind is a messy place in which various thoughts are constantly battling for attention. Introducing a piece of information at just the right moment can provoke people to act upon it.
- **Framing**: The same piece of information is often perceived very differently depending on how it is presented. Framing nudges often make use of this fact to make information more actionable.
- **Simplification and increased convenience**: Complexity frustrates and demotivates. Making something easier increases its chances of being completed.
- **Gaming**: Humans like to play. Curiosity and ludic urges are great motivators for participation and making efforts.
- **Commitment**: People sometimes fall short of their own intentions. Creating a commitment beforehand increases the chance that they will follow through.

Table 4.1 provides some examples of how these common approaches could be used within an integrity system. However, the application of behavioural insights in specific interventions needs to be adapted to the context and respond to the individual behavioural problem. The list presented in Table 4.1 is thus just a starting point for making integrity policies more efficient, more effective and more human.

Table 4.1. Use of tools for choice architecture change in an integrity system

Tool	Typical application	Example
Default	Setting default in procedures	At the beginning of a new project each team member automatically receives a form to declare any conflict of interest
Reminders and framing	Moral messaging	"Remember, you should report and surrender the gifts you receive throughout the year. Show your honesty! Report your gift here."
	Personalised messaging	"Dear Tina, do not forget to take the ethics survey."
	Identity prime	"As a good police officer, you are committed to integrity…"
	Value prime	"Receiving gifts compromises your impartiality as a public official."
	Social prime	"96% of your colleagues have already submitted their asset declarations…"
Simplification and convenience	Reducing the number of steps in a process	Submission of asset declaration through the electronic system
	Chunking (remembering no more than 7 items at a time)	Group the paragraphs of the code of ethics
	Visualisation	Visual comparability of tenders in a graphic
	Colour priming	Use of a red button to submit a sensitive decision
	Convenience	Asking for the asset declaration at the same time as the tax filings are due.
Gaming	Lottery	"You have been selected…
	Games, quizzes	… to participate in an ethics quiz on the elements of the Code of Ethics…
	Competitions	… players from each division compete…
	Micro-incentives	… and the winning division receives a delicious breakfast at the office"
Commitment device	Self-commitment	"I hereby declare that all information entered below in this online recruitment form will be truthful"
	Actionable dilemma training	Considering ethical dilemmas in advance and committing beforehand to handling them well (e.g. preparing a diplomatic response to a dinner invitation that one should not accept).

Source: Inspired by (Sunstein, 2014[112]), as well as (European Commission, 2016[2]); (OECD, 2017[111]); (Van Bavel et al.,(n.d.)[113]).

References

Abbink, K. et al. (2014), "Letting the briber go free: An experiment on mitigating harassment bribes", *Journal of Public Economics*, Vol. 111, pp. 17-28, http://dx.doi.org/10.1016/j.jpubeco.2013.12.012. [100]

Akerlof, G. and J. Yellen (1990), "The fair wage-effort hypothesis and unemployment", *Quarterly Journal of Economics*, Vol. 105/2, pp. 225-283, http://dx.doi.org/10.2307/2937787. [87]

Akerlof, G. and R. Kranton (2011), *Identity Economics*, Hanser, München. [68]

Apesteguia, J., M. Dufwenberg and R. Selten (2007), "Blowing the whistle", *Economic Theory*, Vol. 31/1, pp. 143-166. [101]

Aquino, K. and A. Reed (2002), "The self-importance of moral identity", *Journal of Personality and Social Psychology*, Vol. 83/6, pp. 1423-1440, http://dx.doi.org/10.1037//0022-3514.83.6.1423. [19]

Asch, E. (1951), *Effects of group pressure upon the modification and distortion of judgments.*, Carnegie Press, http://psycnet.apa.org/psycinfo/1952-00803-001 (accessed on 05 July 2017). [99]

Asch, S. (1952), "Group forces in the modification and distortion of judgments", in *Social psychology*, Prentice-Hall, Inc, http://dx.doi.org/10.1037/10025-016. [63]

Ashforth, B. and V. Anand (2003), "The normalization of corruption in organizations", *Research in Organizational Behavior*, Vol. 25, pp. 1-52, http://dx.doi.org/10.1016/S0191-3085(03)25001-2. [119]

Babcock, L. et al. (1995), "Biased judgments of fairness in bargaining", *American Economic Review*, Vol. 85/5, http://www.jstor.org/stable/2950993, pp. 1337-1343. [35]

Babcock, L., X. Wang and G. Loewenstein (1996), "Choosing the wrong pond: Social comparisons in negotiations that reflect a self-serving bias", *Quarterly Journal of Economics*, Vol. 111/1, http://www.jstor.org/stable/2946655 (accessed on 31 May 2017), pp. 1-19. [36]

Banerji, R. (2016), *India has A 'zero rupee note' and it is to be given to corrupt officials who ask you for a bribe* -, Indiatimes.com, http://www.indiatimes.com/culture/who-we-are/india-has-a-zero-rupee-note-and-it-is-to-be-given-to-corrupt-officials-who-ask-you-for-a-bribe-252089.html (accessed on 23 January 2018). [82]

Barkan, R., S. Ayal and D. Ariely (2015), "Ethical dissonance, justifications, and moral behavior", *Current Opinion in Psychology*, Vol. 6, pp. 157-161, http://dx.doi.org/10.1016/j.copsyc.2015.08.001. [15]

Batson, C. et al. (1999), "Moral hypocrisy: Appearing moral to oneself without being so.", *Journal of Personality and Social Psychology*, Vol. 77/3, pp. 525-537, http://dx.doi.org/10.1037/0022-3514.77.3.525. [46]

Bazerman, M., D. Moore and G. Loewenstein (2002), "Why good accountants do bad audits", *Harvard Business Review*, Vol. 80/11, pp. 96-102, https://hbr.org/2002/11/why-good-accountants-do-bad-audits. [105]

Bazerman, M. and A. Tenbrunsel (2011), *Blind spots: Why we fail to do what's right and what to do about it*, Princeton University Press, Princeton. [20]

Benjamin, D., J. Choi and A. Strickland (2010), "Social identity and preferences", *American Economic Review*, Vol. 100/4, pp. 1913-1928, http://www.jstor.org/stable/27871281. [86]

Bicchieri, C. and E. Xiao (2009), "Do the right thing: But only if others do so", *Journal of Behavioral Decision Making*, Vol. 22/2, pp. 191-208, http://dx.doi.org/10.1002/bdm.621. [120]

Boddy, C., R. Ladyshewsky and P. Galvin (2010), "Leaders without ethics in global business: Corporate psychopaths", *Journal of Public Affairs*, Vol. 10/3, pp. 121-138, http://dx.doi.org/10.1002/pa.352. [17]

Bowles, S. and H. Gintis (2011), *A cooperative species: Human reciprocity and its evolution*, Princeton University Press, Princeton. [51]

Buccirossi, P. and G. Spagnolo (2006), "Leniency policies and illegal transactions", *Journal of Public Economics*, Vol. 90/6-7, pp. 1281-1297, http://dx.doi.org/10.1016/j.jpubeco.2005.09.008. [72]

Bursztyn, L. et al. (2016), "Moral incentives in credit card debt repayment: Evidence from a field experiment", *NBER Working Paper, No. 21611*, http://home.uchicago.edu/~bursztyn/Moral_Incentives_20161115.pdf. [25]

Cain, D., G. Loewenstein and D. Moore (2005), "The dirt on coming clean: Perverse effects of disclosing conflicts of interest", *Journal of Legal Studies*, Vol. 34/1, pp. 1-25, http://dx.doi.org/10.1086/426699. [48]

Cain, D., G. Loewenstein and D. Moore (2011), "When sunlight fails to disinfect: Understanding the perverse effects of disclosing conflicts of interest", *Journal of Consumer Research*, Vol. 37/5, pp. 836-857, http://dx.doi.org/10.1086/656252. [47]

Camerer, C. (2003), *Behavioral game theory: Experiments in strategic interaction*, Princeton University Press. [52]

Carlsmith, J. and A. Gross (1969), "Some effects of guilt on compliance", *Journal of Personality and Social Psychology*, Vol. 11/3, pp. 232-239, http://faculty.uncfsu.edu/tvancantfort/Syllabi/Gresearch/Readings/16Carlsmith.pdf. [9]

Chugh, D., M. Bazerman and M. Banaji (2005), *Bounded ethicality as a psychological barrier to recognizing conflicts of interest*, Cambridge University Press, https://sites.fas.harvard.edu/~mrbworks/articles/2005_Chugh_DAMoore.pdf. [34]

Dana, J., R. Weber and J. Kuang (2007), "Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness", *Economic Theory*, Vol. 33/1, pp. 67-80, http://dx.doi.org/10.1007/s00199-006-0153-z. [38]

Derfler-Rozin, R., M. Pillutla and S. Thau (2010), "Social reconnection revisited: The effects of social exclusion risk on reciprocity, trust, and general risk-taking", *Organizational Behavior and Human Decision Processes*, Vol. 112/2, pp. 140-150, http://dx.doi.org/10.1016/j.obhdp.2010.02.005. [103]

Economist (2010), *Taking away Dell's cookie jar*, The Economist Online, https://www.economist.com/blogs/newsbook/2010/07/dells_sec_settlement (accessed on 05 July 2017). [37]

Economist (2017), *Nudge comes to shove: Policymakers around the world are embracing behavioural science*, The Economist Online, http://www.economist.com/news/international/21722163-experimental-iterative-data-driven-approach-gaining-ground-policymakers-around?frsc=dg|a (accessed on 29 May 2017). [1]

Eil, D. and J. Rao (2011), "The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself", *Source: American Economic Journal: Microeconomics American Economie Journal: Microeconomics*, Vol. 3/3, pp. 114-138, http://www.jstor.org/stable/41237187 (accessed on 06 June 2017). [84]

Erat, S. (2013), "Avoiding lying: The case of delegated deception", *Journal of Economic Behavior and Organization*, Vol. 93, pp. 273-278, http://dx.doi.org/10.1016/j.jebo.2013.03.035. [118]

European Commission (2016), *Behavioural Insights Applied to Policy. European Report 2016*, European Union, http://dx.doi.org/10.2760/903938. [2]

Falk, A. and M. Kosfeld (2006), "The hidden costs of control", *American Economic Review*, Vol. 96/5, pp. 1611-1630, http://www.jstor.org/stable/30034987. [44]

Falk, A. and N. Szech (2013), "Organizations, diffused pivotality and immoral outcomes", https://polit.econ.kit.edu/downloads/papers/Falk_Szech_WP_2_Pivotality_Falk_Szech_DP-2.pdf (accessed on 03 October 2017). [41]

Fehr, E. and S. Gächter (2000), "Fairness and retaliation: The economics of reciprocity", *Journal of Economic Perspectives*, Vol. 14/3, pp. 159-181, http://www.jstor.org/stable/2646924. [54]

Fehr, E. and J. List (2004), "The hidden costs and return of incentives: Trust and trustworthiness among CEOs", *Journal of the European Economic Association*, Vol. 2/5, pp. 743-771, http://dx.doi.org/10.1162/1542476042782297. [93]

Fellner, G., R. Sausgruber and C. Traxler (2013), "Testing enforcement strategies in the field: Threat, moral appeal and social information", *Journal of the European Economic Association*, Vol. 11/3, pp. 634-660, http://dx.doi.org/10.1111/jeea.12013. [29]

Gilchrist, D., M. Luca and D. Malhotra (forthcoming), "When 3+1 is more than 4: Gift structure and reciprocity in the field", *Management Science,*, http://people.hbs.edu/mluca/oDesk.pdf (accessed on 03 August 2017). [88]

Gino, F. and M. Bazerman (2009), "When misconduct goes unnoticed: The acceptability of gradual erosion in others' unethical behavior", *Journal of Experimental Social Psychology*, Vol. 45/4, pp. 708-719, http://dx.doi.org/10.1016/j.jesp.2009.03.013. [66]

Gino, F. et al. (2011), "Unable to resist temptation: How self-control depletion promotes unethical behavior", *Organizational Behavior and Human Decision Processes*, Vol. 115/2, pp. 191-203, http://dx.doi.org/10.1016/j.obhdp.2011.03.001. [14]

Gino, F. and A. Galinsky (2012), "Vicarious dishonesty: When psychological closeness creates distance from one's moral compass", *Organizational Behavior and Human Decision Processes*, Vol. 119/1, pp. 15-26, http://dx.doi.org/10.1016/j.obhdp.2012.03.011. [83]

Gino, F. and C. Mogilner (2014), "Time, Money, and Morality", *Psychological Science*, Vol. 25/2, pp. 414-421, http://dx.doi.org/10.1177/0956797613506438. [28]

Gneezy, U., S. Meier and P. Rey-Biel (2011), "When and why incentives (don't) work to modify behavior", *Journal of Economic Perspectives*, Vol. 25/4, pp. 191-210, http://dx.doi.org/10.1257/jep.25.4.191. [90]

Gneezy, U., A. Imas and K. Madarász (2014), "Conscience accounting: Emotion dynamics and social behavior", *Management Science*, Vol. 60/11, pp. 2645-2658, http://dx.doi.org/10.1287/mnsc.2014.1942. [10]

Godwin, A. (2009), *Ethical decision making*, Office of the Merit Protection Commissioner, http://www.apsc.gov.au/publications-and-media/current-publications/ethical-decision-making (accessed on 08 March 2018). [95]

Hamman, J., G. Loewenstein and R. Weber (2010), "Self-interest through delegation: An additional rationale for the principal-agent relationship", *American Economic Review*, Vol. 100, pp. 1826-1846, http://dx.doi.org/10.1257/aer.100.4.1826. [42]

Hare, R. (2006), "Psychopathy: A clinical and forensic overview", *Psychiatric Clinics of North America*, Vol. 29/3, pp. 709-724, http://dx.doi.org/10.1016/j.psc.2006.04.007. [16]

Henrich, J. et al. (2005), "'Economic man' in cross-cultural perspective: Behavioral experiments in 15 small-scale societies", *Behavioral and Brain Sciences*, Vol. 28/06, pp. 795-815, http://dx.doi.org/10.1017/S0140525X05000142. [53]

Hicken, A. et al. (2015), "Measuring vote-selling: Field evidence from the Philippines", *American Economic Review: Papers & Proceedings*, Vol. 105/5, pp. 352-356, http://dx.doi.org/10.1257/aer.p20151033. [24]

Hildreth, J., F. Gino and M. Bazerman (2016), "Blind loyalty? When group loyalty makes us see evil or engage in it", *Organizational Behavior and Human Decision Processes*, Vol. 132, pp. 16-36, http://dx.doi.org/10.1016/j.obhdp.2015.10.001. [85]

Hoffman, M., E. Yoeli and A. Nowak (2015), "Cooperate without looking: Why we care what people think and not just what they do", *Proceeding of the National Academy of Sciences*, Vol. 112/6, pp. 1727-1732, http://dx.doi.org/10.1073/pnas.1417904112. [98]

Houser, D., S. Vetter and J. Winter (2012), "Fairness and cheating", *European Economic Review*, Vol. 56/8, pp. 1645-1655, http://dx.doi.org/10.1016/j.euroecorev.2012.08.001. [7]

Husted, B. (1994), "Honor among thieves: A transaction-cost interpretation of corruption in third World countries", *Business Ethics Quarterly*, Vol. 4/1, pp. 17-27, http://dx.doi.org/10.2307/3857556. [56]

International Organization of Supreme Audit Institutions (2010), *Internal Audit Independence in the Public Sector, INTOSAI GOV 9140*, International Organization of Supreme Audit Institutions, Vienna. [107]

Kahneman, D. and A. Tversky (1979), "Prospect theory: An analysis of decision under risk", *Econometrica*, Vol. 47/2, pp. 263-292, http://dx.doi.org/10.2307/1914185. [102]

Kalanthroff, E., C. Aslan and R. Dar (2017), "Washing away your sins will set your mind free: Physical cleansing modulates the effect of threatened morality on executive control", *Cognition and Emotion*, Vol. 31/1, pp. 185-192, http://dx.doi.org/10.1080/02699931.2015.1086313. [11]

Kennedy, J. and C. Anderson (2017), "Hierarchical rank and principled dissent: How holding higher rank suppresses objection to unethical practices", *Organizational Behavior and Human Decision Processes*, Vol. 139, pp. 30-49, http://dx.doi.org/10.1016/j.obhdp.2017.01.002. [96]

Kern, M. and D. Chugh (2009), "Bounded ethicality: The perils of loss framing", *Psychological Science*, Vol. 20/3, pp. 378-384, http://dx.doi.org/10.1111/j.1467-9280.2009.02296.x. [110]

Kosfeld, M. and S. Neckermann (2011), "Getting more work for nothing? Symbolic awards and worker performance", *American Economic Journal: Microeconomics*, Vol. 3, pp. 86-99, http://dx.doi.org/10.1257/mic.3.3.86. [92]

Kosfeld, M., S. Neckermann and X. Yang (2014), "Knowing that you matter, matters! The interplay of meaning, monetary incentives, and worker recognition", *SSRN Electronic Journal*, http://dx.doi.org/10.2139/ssrn.2533442. [91]

Kouchaki, M. et al. (2013), "Seeing green: Mere exposure to money triggers a business decision frame and unethical outcomes", *Organizational Behavior and Human Decision Processes*, Vol. 121/1, pp. 53-61, http://dx.doi.org/10.1016/j.obhdp.2012.12.002. [114]

Kouchaki, M. and I. Smith (2014), "The Morning Morality Effect", *Psychological Science*, Vol. 25/1, pp. 95-102, http://dx.doi.org/10.1177/0956797613498099. [21]

Lambsdorff, J. and M. Nell (2007), "Fighting corruption with asymmetric penalties and leniency", *Center for European, Governance and Economic Development Research Discussion Papers*, https://ideas.repec.org/p/zbw/cegedp/59.html (accessed on 30 January 2018). [73]

Lambsdorff, J. (2012), "Behavioural and experimental economics as a guidance to anticorruption", in Serra, D. and L. Wantchekon (eds.), *New Advances in Experimental Research on Corruption*, Emerald. [55]

Lambsdorff, J. (2015), "Preventing corruption by promoting trust – Insights from behavioral science", *Passauer Diskussionspapiere*, No. V-69-15, Universität Passau, Passau, http://dx.doi.org/10.13140/RG.2.1.3563.4006. [45]

Lambsdorff, J. (2016), *A Behavioral science approach to preventing corruption, Global Anticorruption Blog*, https://globalanticorruptionblog.com/2016/02/11/guest-post-a-behavioral-science-approach-to-preventing-corruption/ (accessed on 28 September 2017). [75]

Loewenstein, G. et al. (1993), "Self-serving assessments of fairness and pretrial bargaining", *Journal of Legal Studies*, Vol. 22/1, pp. 135-159, http://dx.doi.org/10.1086/468160. [32]

Loewenstein, G., S. Sah and D. Cain (2012), "The unintended consequences of conflict of interest disclosure", Vol. 307/7, http://bevwin.pbworks.com/w/file/fetch/51097704/JAMA.COI.pdf (accessed on 06 June 2017), pp. 669-670. [49]

Mazar, N. and D. Ariely (2006), "Dishonesty in everyday life and its policy implications", *Journal of Public Policy and Marketing*, Vol. 25/1, pp. 117-126, http://www.jstor.org/stable/30000530. [26]

Mazar, N., O. Amir and D. Ariely (2008), "The dishonesty of honest people: A theory of self-concept maintenance", *Journal of Marketing Research*, Vol. 45/6, pp. 633-644, http://www.jstor.org/stable/20618852. [13]

Milgram, S. (1963), "Behavioral Study of Obedience", *Journal of Abnormal and Social Psychology*, Vol. 67/4, pp. 371-378, http://library.nhsggc.org.uk/mediaAssets/Mental%20Health%20Partnership/Peper%202%2027th%20Nov%20Milgram_Study%20KT.pdf (accessed on 22 January 2018). [64]

Miller, G. (1955), "The magical number seven, plus or minus two: Some limits on our capacity for processing information", *Psychological Review*, Vol. 101/2, pp. 343-352, http://www.psych.utoronto.ca/users/peterson/psy430s2001/Miller%20GA%20Magical%20Seven%20Psych%20Review%201955.pdf (accessed on 24 January 2018). [78]

Ministry of Finance (2015), *Seven key duties for civil servants in central government*, http://www.modst.dk/media/17483/kodex_vii_english_version.pdf (accessed on 24 January 2018). [79]

Moore, C. and F. Gino (2013), "Ethically adrift: How others pull our moral compass from true North, and how we can fix it", *Research in Organizational Behavior*, Vol. 33, pp. 53-77, http://dx.doi.org/10.1016/j.riob.2013.08.001. [58]

Nowak, M. and K. Sigmund (2005), "Evolution of indirect reciprocity", *Nature*, Vol. 437/7063, pp. 1291-1298, http://dx.doi.org/10.1038/nature04131. [61]

OECD (2009), *Towards a Sound Integrity Framework: Instruments, Processes, Structures and Conditions for Implementation (GOV/PGC/GF(2009)1)*. [94]

OECD (2016), *Being an Independent Regulator*, The Governance of Regulators, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264255401-en. [108]

OECD (2017), *Behavioural Insights and Public Policy: Lessons from Around the World*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264270480-en. [3]

OECD (2017), *OECD Recommendation of the Council on Public Integrity*, http://www.oecd.org/gov/ethics/recommendation-public-integrity.htm. [5]

OECD (2017), *Creating a Culture of Independence: Practical Guidance against Undue Influence*, The Governance of Regulators, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264274198-en. [30]

OECD (2017), *Preventing Policy Capture: Integrity in Public Decision Making*, OECD Public Governance Reviews, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264065239-en. [31]

OECD (2017), *Tackling Environmental Problems with the Help of Behavioural Insights*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264273887-en. [111]

Ordnez, L. and D. Welsh (2015), *Immoral goals: How goal setting may lead to unethical behavior*, http://dx.doi.org/10.1016/j.copsyc.2015.06.001. [76]

Ostrom, E. (2000), "Collective action and the evolution of social norms", *Journal of Economic Perspectives*, Vol. 14/3, pp. 137-158, http://dx.doi.org/10.1080/19390459.2014.935173. [71]

Paharia, N. et al. (2009), "Dirty work, clean hands: The moral psychology of indirect agency", *Organizational Behaviour and Human Decision*, Vol. 109/2, pp. 134-141, http://dx.doi.org/10.1016/j.obhdp.2009.03.002. [117]

Persson, A., B. Rothstein and J. Teorell (2013), "Why anticorruption reforms fail: Systemic corruption as a collective action problem", *Governance*, Vol. 26/3, pp. 449-471, http://dx.doi.org/10.1111/j.1468-0491.2012.01604.x. [74]

Pillutla, M. and J. Murnighan (1995), "Being fair or appearing fair: Strategic behavior in ultimatum bargaining", *Academy of Management Journal*, Vol. 38/5, pp. 1408-1426, http://dx.doi.org/10.2307/256863. [39]

Pittarello, A. et al. (2015), "Justifications shape ethical blind spots", *Psychological Science*, Vol. 26/6, pp. 709-804, http://dx.doi.org/10.1177/0956797615571018. [116]

Ploner, M. and T. Regner (2013), "Self-image and moral balancing: An experimental analysis", *Journal of Economic Behaviors and Organization*, Vol. 93, pp. 374-383, http://dx.doi.org/10.1016/j.jebo.2013.03.030. [8]

Pruckner, G. and R. Sausgruber (2013), "Honesty on the streets: A field study on newspaper purchasing", *Journal of the European Economic Association*, Vol. 11/3, pp. 661-679, http://dx.doi.org/10.1111/jeea.12016. [27]

Rand, D. et al. (2009), "Positive interactions promote public cooperation.", *Science*, Vol. 325/5945, pp. 1272-1275, http://dx.doi.org/10.1126/science.1177418. [70]

Reuben, E. and M. Stephenson (2013), "Nobody likes a rat: On the willingness to report lies and the consequences thereof", *Journal of Economic Behavior and Organization*, Vol. 93, pp. 384-391, http://dx.doi.org/10.1016/j.jebo.2013.03.028. [97]

Robert, I. and M. Arnab (2013), "Is dishonesty contagious?", *Economic Inquiry*, Vol. 51/1, pp. 722-734, http://dx.doi.org/10.1111/j.1465-7295.2012.00470.x. [67]

Sachdeva, S., R. Iliev and D. Medin (2009), "Sinning saints and saintly sinners: The paradox of moral self-regulation", *Psychological Science*, Vol. 20/4, pp. 523-528, http://dx.doi.org/10.1111/j.1467-9280.2009.02326.x. [6]

Sah, S., G. Loewenstein and D. Cain (2013), "The burden of disclosure: Increased compliance with distrusted advice", *Journal of Personality and Social Psychology*, Vol. 104/2, pp. 289-304, http://psycnet.apa.org/index.cfm?fa=buy.optionToBuy&id=2012-28199-001 (accessed on 06 June 2017). [50]

Schikora, J. (2011), "Bringing the four-eyes-principle to the lab", *Münchener Wirtschaftswissenschaftliche Beiträge: VWL: discussion papers*, http://www.econbiz.de/Record/bringing-the-four-eyes-principle-to-the-lab-schikora-jan-theodor/10008858142 (accessed on 16 May 2017). [60]

Schulze, G. and B. Frank (2003), "Deterrence versus intrinsic motivation: Experimental evidence on the determinants of corruptibility", *Economics of Governance*, Vol. 4, pp. 143-160, http://dx.doi.org/10.1007/s101010200059. [43]

Sezer, O., F. Gino and M. Bazerman (2015), "Ethical blind spots: Explaining unintentional unethical behavior", *Current Opinion in Psychology*, Vol. 6, pp. 77-81, http://dx.doi.org/10.1016/j.copsyc.2015.03.030. [106]

Shalvi, S. et al. (2011), "Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior", *Organizational Behavior and Human Decision Processes*, Vol. 115/2, pp. 181-190, http://dx.doi.org/10.1016/j.obhdp.2011.02.001. [40]

Shalvi, S., O. Eldar and Y. Bereby-Meyer (2012), "Honesty requires time (and lack of justifications)", *Psychological Science*, Vol. 23/10, pp. 1264-1270, http://dx.doi.org/10.1177/0956797612443835. [22]

Shalvi, S. et al. (2015), "Self-serving justifications: Doing wrong and feeling moral", *Current Directions in Psychological Science*, Vol. 24/2, pp. 125-130, http://dx.doi.org/10.1177/0963721414553264. [57]

Sheldon, O. and A. Fishbach (2015), "Anticipating and resisting the temptation to behave unethically", *Personality and Social Psychology Bulletin*, Vol. 41/7, pp. 962-975, http://dx.doi.org/10.1177/0146167215586196. [23]

Shu, L. et al. (2012), "Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end.", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 109/38, pp. 15197-15200, http://dx.doi.org/10.1073/pnas.1209746109. [80]

Stevens, G., J. Deuling and A. Armenakis (2012), "Successful psychopaths: Are they unethical decision-makers and why?", *Journal of Business Ethics*, Vol. 105/2, pp. 139-149, http://dx.doi.org/10.1007/s10551-011-0963-1. [18]

Sunstein, C. (2014), "Nudging: A very short guide", *Journal of Consumer Policy*, Vol. 37/4, pp. 583-588, http://dx.doi.org/10.1007/s10603-014-9273-1. [112]

Tarrant, M. et al. (2012), "Social identity and perceptions of torture: It's moral when we do it", *Journal of Experimental Social Psychology*, Vol. 48/2, pp. 513-518, http://dx.doi.org/10.1016/j.jesp.2011.10.017. [33]

Tenbrunsel, A. (1998), "Misrepresentation and expectations of misrepresentation in an ethical dilemma: The role of incentives and temptation", *Academy of Management Journal Academy of Management Journal*, Vol. 41/3, pp. 330-339, http://www.jstor.org/stable/256911 (accessed on 04 July 2017). [115]

Thau, S. et al. (2015), "Unethical for the sake of the group: Risk of social exclusion and pro-group unethical behavior", *Journal of Applied Psychology*, Vol. 100/1, pp. 98-113, http://ink.library.smu.edu.sg/lkcsb_research (accessed on 03 August 2017). [104]

Tversky, A. and D. Kahneman (1973), "Availability: A heuristic for judging frequency and probability", *Cognitive Psychology*, Vol. 5/2, pp. 207-232, http://dx.doi.org/10.1016/0010-0285(73)90033-9. [109]

Tyler, T. (1999), "Why people cooperate with organizations: An identity-based perspective", *Research on Organizational Justice*, Vol. 21, pp. 201-246. [69]

Van Bavel, R. et al. (n.d.), "Applying Behavioural Sciences to EU Policy-making", http://dx.doi.org/10.2788/4659. [113]

Van Veldhuizen, R. (2013), "The influence of wages on public officials' corruptibility: A laboratory investigation", *Journal of Economic Psychology*, Vol. 39, pp. 341-356, http://dx.doi.org/10.1016/j.joep.2013.09.009. [89]

Welsh, D. and L. Ordóñez (2014), "The dark side of consecutive high performance goals: Linking goal setting, depletion, and unethical behavior", *Organizational Behavior and Human Decision Processes*, Vol. 123/2, pp. 79-89, http://dx.doi.org/10.1016/j.obhdp.2013.07.006. [77]

Wiltermuth, S. (2011), "Cheating more when the spoils are split", *Organizational Behavior and Human Decision Processes*, Vol. 115/2, pp. 157-168, http://dx.doi.org/10.1016/j.obhdp.2010.10.001. [59]

World Bank (2015), *Mind, Society, and Behavior*, World Bank, Washington, DC, http://documents.worldbank.org/curated/en/645741468339541646/pdf/928630WDR0978100Box385358B00PUBLIC0.pdf (accessed on 24 November 2017). [4]

Yoeli, E. et al. (2013), "Powering up with indirect reciprocity in a large-scale field experiment", *Proceedings of the National Academy of Sciences*, Vol. 110/Supplement 2, pp. 10424-10429, http://dx.doi.org/10.1073/pnas.1301210110. [62]

Zhong, C. and K. Liljenquist (2006), "Washing away your sins: Threatened morality and physical cleansing", *Science*, Vol. 313/5792, pp. 1451-1452, http://dx.doi.org/10.1126/science.1130726. [12]

Zimbardo, P. and A. Cross (1971), *The Stanford Prison Experiment: A Simulation Study of the Psychology of Imprisonment Conducted August 1971 at Stanford University*, Austin ABC, https://books.google.fr/books?id=cHXwngEACAAJ. [65]

Zinnbauer, D. (2012), "'Ambient accountability' - Fighting corruption when and where it happens", *SSRN Electronic Journal*, http://dx.doi.org/10.2139/ssrn.2168063. [81]

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where governments work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, Chile, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The European Union takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

OECD PUBLISHING, 2, rue André-Pascal, 75775 PARIS CEDEX 16
(04 2018 04 1 P) ISBN 978-92-64-29705-0 – 2018

OECD Public Governance Reviews

Behavioural Insights for Public Integrity

HARNESSING THE HUMAN FACTOR TO COUNTER CORRUPTION

This series includes international studies and country-specific reviews of government efforts to make the public sector more efficient, effective, innovative and responsive to citizens' needs and expectations. Publications in this series look at topics such as open government, preventing corruption and promoting integrity in the public service, risk management, illicit trade, audit institutions, and civil service reform. Country-specific reviews assess a public administration's ability to achieve government objectives and preparedness to address current and future challenges. In analysing how a country's public administration works, reviews focus on cross-departmental co-operation, the relationships between levels of government and with citizens and businesses, innovation and quality of public services, and the impact of information technology on the work of government and its interaction with businesses and citizens.

At the core of integrity is a human choice. Behavioural research reveals two determinants of integrity: the internal dynamics of how individuals make moral decisions and how these decisions are shaped and influenced by other people. This report describes how behavioural insights can help make integrity and anti-corruption policies more effective and efficient. It includes concrete policy applications and provides guidance for policy makers on how to use behavioural insights in designing integrity and anti-corruption policies.

Consult this publication on line at ***http://dx.doi.org/10.1787/9789264297067-en***.

This work is published on the OECD iLibrary, which gathers all OECD books, periodicals and statistical databases. Visit ***www.oecd-ilibrary.org*** for more information.

OECD*publishing*
www.oecd.org/publishing

THE LONDON BOOK FAIR WINNER
INTERNATIONAL EXCELLENCE Awards 2017
IN PARTNERSHIP WITH THE PUBLISHERS ASSOCIATION

ISBN 978-92-64-29705-0
04 2018 04 1 P

9 789264 297050