The Law of Good People

CHALLENGING STATES' ABILITY TO REGULATE HUMAN BEHAVIOR

YUVAL FELDMAN

Bar-Ilan University

CAMBRIDGE
UNIVERSITY PRESS

Contents

Preface

Plato has famously argued that "laws are made to instruct the good, and in the hope that there may be no need of them; also to control the bad, whose hardness of heart will not be hindered from crime."[1]

The premise of this book is that "good" people need laws, but with a different regulatory approach. For law to be able to operate effectively and lead to behavioral change, it needs to understand why "good" people engage in wrongdoing. Yet, even today, the law's main purpose is understood as protecting others from the actions of "bad" people. In contrast, this book argues that it is "good" people (more accurately, those who think of themselves as good people) whose behavior should occupy more of the attention of the law. While, for understandable reasons, the legal system is primarily focused on the upper tail of the distribution of misconduct, in fact most misconduct – those acts we refer to in the book as ordinary unethicality – is actually found in the middle of that curve; in many contexts, its treatment requires completely different assumptions and consequently a modified set of regulatory and enforcement tools from those used to punish serious misconduct.

This book assumes that ordinary unethicality is the most common and relevant type of behavior that legal policy makers should attempt to regulate. Acts of ordinary unethicality are part of most private law disputes in areas such as contact breach, tortious behavior, and lack of respect for people's property rights.[2] In addition, much of people's misconduct in public law, such as tax law, administrative law, and corporate law, involve ordinary unethicality.

Research into behavioral ethics, a growing field within the psychology and management literatures, has demonstrated in numerous field and lab experiments that people's unethical behavior stems from numerous cognitive and social processes that are only partially related to people's deliberative and aware reasoning;

---

[1] Taken from Jowett, B. (1871). *The dialogues of Plato* (Vol. 8). Oxford: Oxford University Press, p. 128. Interestingly, the following "Good people do not need laws to tell them to act responsibly, while bad people will find a way around the laws" is a misquoted phrase of Plato that is indeed better for the purposes of the book.

[2] Both tangible property and even more so with regard to intellectual property.

hence, people can engage in misconduct and still think of themselves as "good" people.

For the purposes of the book, we divide good people into "authentic good people," who through various implicit self-serving processes, such as blind spots or moral forgetting, understand reality in a biased way without recognizing it, and situational wrongdoers who use various justifications in the situation to behave badly without feeling immoral.[3] In contrast to the current dominant enforcement paradigm, which assumes that states are dealing with bad people who are pursuing their own self-interest and that the law needs to increase the price for misbehavior through sanctions and punishment, behavioral ethics posits that good people are guided by both automatic and deliberative processes and biases that enable them to bend the laws within the confines of their conscience. The book claims that the regulatory treatment of the good people who commit many of these misconducts is the neglected task of the law and of legal theory. A focus on those good people will enable regulators to determine in advance which types of situations are likely to encourage acts of ordinary unethicality and to develop an effective regulatory toolbox.

Although behavioral approaches have been introduced into legal theory and practice in the past two decades, they have been narrowly focused on the biases and heuristics literature, which has served as the foundation of the field of behavioral economics: that dominant field deals with cognitive biases in decision making in contexts related to probability, risk, and money. In contrast, behavioral ethics, which focuses on people's biases in making ethical decisions, has received little attention. Yet, ethical biases are not only stronger in many ways than the cognitive biases of behavioral economics but also far more relevant for the regulation of people's ethical behavior. For example, whereas behavioral economics uses the optimism bias to understand why people pursue legal disputes in courts, the behavioral ethics approach suggests that people go to court because they fail to fully understand how problematic their own behavior is, both legally and morally; various self-maintenance mechanisms prevent them from recognizing their wrongdoing. In addition, loss aversion affects not just risk perception but also the likelihood that people will behave unethically.

To regulate the involvement of "good" people in various types of noncooperative behaviors, this book provides a broad theoretical and empirical comparison of traditional and nontraditional enforcement mechanisms. It argues that many of the existing regulatory enforcement strategies are not suitable for addressing misconduct stemming from both deliberative and non-deliberative reasoning processes and biases. The insights of behavioral ethics into the cognitive and motivational factors guiding the behavior of good people require the development of innovative

---

[3] Obviously the majority of good people are found on a spectrum between those prototypes of good people, using both deliberative and implicit mechanisms to justify their unethically.

approaches to the normative treatment of a diverse population consisting of both good and bad people. These new approaches use an *ex ante* approach that attempts to regulate behavior in advance rather than an *ex post* approach that tries to determine responsibility, which behavioral ethics suggests is quite complex. The move to an *ex ante* design reduces the need to find the smoking guns required by evidentiary rules; hence, it gives more importance to dealing with minor misconducts that cannot justify *ex post* legal examination but in the aggregate are not less important. Good people's ignorance and unawareness of the moral meaning of their behavior provide another justification for a greater focus on detection than on punishment. This new paradigm facilitates a more nuanced understanding by regulators of the likely implicit and explicit effects on behaviors of specific situations. Finally, it suggests that differentiated approaches focusing on variation in people's level of intrinsic motivation should be supplemented by strategies that account for variation in people's level of self-awareness of the moral and legal meaning of their behavior.

The book examines the law's effectiveness in preventing people from engaging in uncooperative behaviors and wrongful conducts, such as breaching contracts, engaging in corruption and employment discrimination, and eschewing professional duties. More specifically, it compares the impact of traditional methods, including deterrence, social norms, and procedural justice, with that of behaviorally informed enforcement mechanisms, such as nudges, framing, and debiasing. The book discusses the pros and cons of these various intervention mechanisms, drawing practical conclusions for legal policy makers on how to optimize their regulatory and enforcement effects on both the deliberative and non-deliberative components of unethical behavior.

In addition, this book addresses key unresolved theoretical questions from several directions. How much can we know ex ante about good people's awareness and ability to control their unethical behaviors? How can we know that their goodness is genuine and not faked? Are considerations of morality and traditional enforcement practices, such as deterrence, effective in curbing behaviors that are only partly deliberative? Can states regulate simultaneously different types of individuals by using different intervention methods? Should the nudge approach, which avoids direct communication between the state and the people it regulates, replace all other intervention methods? Do we know what is lost in the sustainability of behavioral change and in autonomy when we abandon traditional intervention methods and replace them with interventions that don't require any deliberative reasoning by people? Is there still a benefit derived from changing people's intrinsic motivation when many of their misconducts are not done with full awareness? The book addresses these questions and examines in what way the existing research falls short of offering a coherent behavioral and normative picture of the person we are trying to regulate.

In its call for regulatory reform, the book does not focus only on theoretical discussions; instead, it draws on extensive empirical research that other researchers and I have conducted on these questions. It examines through case studies, the effect of social norms on the perception of legality in the context of intellectual property, the effect of legal incentives on people's intrinsic and extrinsic motivations in the area of environmental protection and whistle-blowing laws, the effect of legal uncertainty on the compliance and performance of people with different motivational backgrounds, and the effectiveness of deterrence and morality in enabling people to avoid subtle conflicts of interest and refrain from engaging in implicit discrimination toward different social groups.

Chapters 1 and 2 lay the groundwork for this innovative approach to the law. Chapter 1 outlines the main argument of the book, briefly describing the potential of behavioral ethics, exposing existing gaps in the behavioral analysis of law, and showing how the book proposes to fill them. Chapter 2 explains how both deliberate mechanisms, (e.g. moral disengagement) and non-deliberate mechanisms, (e.g. moral forgetting and motivated blindness) prevent people from recognizing the wrongdoing in their behavior and their own unethicality.

Chapters 3–5 address how to expand the regulatory toolbox, focusing on both formal and non-formal controls and their ability to deal with both explicit and implicit types of misconducts by both "good" and "bad" people. These chapters outline factors such as situational design, behavioral incentives, social and ethical nudges, fairness, social norms, and education. The focus on good people requires a shift in the focus of the legal regime from *ex post* liability to *ex ante* design. *Ex post* mechanisms that focus on liability and are designed to change people's *ex ante* calculations will not be effective because most "good" people are not likely to be aware of why they behaved in a certain way in the first place. We also argue for the importance of designing policies that make it difficult for people to interpret fairness in a self-serving way. There is a need to provide people with accurate information on the nature of social norms and their prevalence, because various cognitive mechanisms are likely to cause people to underestimate the true prevalence of cooperative norms. Incentives need to be sensitive to the crowding-out effect of intrinsic motivation and hence should account for people's motivational sensitivities. Ethical nudges need to be distinguished from other kinds of behavioral nudges, so that appeals to self-interest do not reduce their effectiveness.

Chapter 6, which focuses on individual differences, complicates the picture of good vs. bad people by arguing that good people should be divided into at least two types[4], based on the variation in self-deception mechanisms reviewed in Chapter 2. Good people of the first type genuinely do not understand their behavior or the relevant situation as it is, because of cognitive mechanisms, such as Bazerman's blind spot, Haidt's "emotional dog" approach to morality, or Balcetis's motivated

---

[4] This dichotomy is first developed in Chapter 4.

seeing, which rely mostly on non-deliberative processes. In that camp of morally blind people belong those who engage in implicit job discrimination or in implicit corruption in subtle conflict of interest situations. In contrast, good people of the second type know that what they are doing is impermissible, but they find various rationales that allow them to do bad things without harming their moral self-image. The related work of Bandura on moral disengagement and Shalvi's work on justified dishonesty in the dice-under-the-cup paradigm shows that this is mostly a deliberate process. In the camp of the deliberative justifiers, we can find those committing various parking violations, cutting corners, or using personal contacts to bypass a certain bureaucratic procedure.[5] In this chapter, we also analyze relevant individual differences scales, such as moral identity and social value orientation. Recognizing the limitations of identifying individual variation ex ante, the chapter concludes with a discussion of alternative ways to differentiate between different types of people based on their commitment to the law or the norms of their profession. Such approaches are more likely to be known ex ante by regulators.

The complex effects of law, the existence of various types of regulatory tools, the power of the situation, and the variation among people together create a highly complex regulatory picture. Both Chapter 7 on the pluralistic account of law and Chapter 8 on the trade-offs between the different likely effects of laws present research that guides how policy makers can balance the effects of different aspects of the law, on different people, with regard to different compliance behaviors. The concept of the pluralistic approach to law refers to the fact that the law tries to change the behavior of people with different compliance motivations and different levels of awareness, and through the use of both implicit and explicit signals; thus, there are many possible conflicting behavioral outcomes to any law that is enacted. The concept of behavioral trade-offs suggests that each type of intervention produces a different behavioral reaction, because the behavioral ethics perspective on people's approach to unethicality is far more complex than is assumed by traditional enforcement methods. It is rarely the case that any one policy will be superior in terms of every behavioral dimension. Developing an effective policy requires taking many factors into consideration, as well as the particular context of the situations in which people make decisions.

This book's approach to law enforcement is applied in Chapters 9 and 10 to corruption and employment discrimination, respectively; these two case studies demonstrate how to create an effective balance between regulatory tools that address different types of populations with different mind-sets toward the behavior to be regulated. For both corruption and employment discrimination, it is not the behavior itself, but the state of mind of the individual when doing it that is problematic – and that is very hard to prove in court. For example, when a mayor hires a contractor,

---

[5] As suggested above, many good people are on a spectrum between those two prototypes, depending on the specific situation.

that behavior is only of concern when that decision is made based at least partially on his or her self-interest. When an employer decides not to hire someone for a job, that decision is problematic if he or she based it on prejudice. These chapters examine the different set of regulatory tools to be used *ex ante* rather than *ex post* to deal with legal violations in these two fields.

Chapter 11 concludes with a discussion of some key policy-making concepts derived from the book's assumptions about wrongdoing and the good vs. bad people typology. It first examines the role of intrinsic motivation and moral education when decisions are non-deliberative. It then outlines the use of taxonomies of regulatory contexts to determine which regulatory tools to use in which contexts and for which purposes and reviews possible approaches to deal with people's unethical and illegal behavior, such as differentiated or targeted regulation and responsive or sequential regulation. It concludes with a discussion of the future of the field of behavioral ethics in the theory of legal enforcement and some of its limitations.

Acknowledgments

lab headed by Mahzarin Banaji, also at Harvard University. Most of my knowledge about behavioral ethics was attained during this time. The intellectual tools I gained while working in these labs have exerted a considerable influence on the course and direction of my research. The creation of this book evinces this influence, for which I am deeply grateful.

Some of the research in this book can also be traced to an even earlier – my doctoral studies at the University of California–Berkeley. Naturally, my advisors Bob Cooter and Rob MacCoun – whose different but equally sophisticated and coherent views on rationality and decision-making processes helped shape my current view about "good people" – require much gratitude. I would also like to thank the entire Jurisprudence and Social Policy faculty at Boalt Hall for their indirect help with the creation of this book. It took me many years to recognize the importance of studying not just the behavioral and economic approaches to law but also the sociological, philosophical, and political science views on decision making in legal contexts.

The vast majority of my research has been conducted in collaboration with scholars from numerous disciplines, such as law, psychology, sociology, political science, philosophy, economics, and game theory. Generally, credit is given to each person in the appropriate chapter.

However, since so much of my research has emerged from that collaborative work, I feel the need to thank each one of them for their specific contribution to this work as a whole:

To Doron Teichman, for a series of joint papers regarding the nuanced effect of law on behavior, the last three being in collaboration with Amos Schur and Uriel Haran.

To Henry Smith, Constantine Boassalis, and Alon Harel, for our joint work on theoretical and empirical aspects of legal ambiguity.

To Orly Lobel, for our joint work on legal instruments, behavioral trade-offs, whistle-blowing, and new regulatory tools.

To Oren Perez, for our joint work on various legal incentives and their effect on environmental behavior.

To Tom Tyler, for our joint work on the effect of law on procedural justice.

To Eliran Halali, for our joint work on competing approaches to enforcing behavior in subtle conflict of interest situations.

To Michal Feldman and Bob Cooter, for our joint work on biases in the perception of norms.

To Tammi Krichelli-Katz and Haggay Porat, for our joint work regarding employment discrimination.

To Maryam Koachaki and Francesca Gino, for our joint work regarding the expressive effects of ethical codes.

Finally, I would like to extend my gratitude to the numerous research assistants who helped in various stages of this book's preparation: Noam Loshinsky, Jasmin Goldofsky, Russel Shitrit-Leibovitch, Or Ashual, Aelet Sender, Maryam Jacobi,

1

Introduction

The focus of this book is on how governments may effectively use recent advances in the understanding of human behavior to guide their efforts to modify people's behavior. To date, the insights of behavioral ethics that have completely revolutionized the business and management fields have yet to be applied in legal theory and policy research, especially in the context of legal enforcement and compliance. The growing recognition that misconduct can be facilitated by structural issues and is not just the product of a few "bad apples" has important implications for the creation and fine-tuning of institutional design and enforcement mechanisms. States need to modify their regulatory roles and functions based on the understanding that discrimination does not just stem from certain employers who hate minorities, that corruption is not just about greedy individuals, or that trade secrets are not just divulged for mercenary motives.

This book argues that the good-people rationale – the idea that ordinary people could engage in all types of wrongdoing without being aware of the full meaning of their behavior – greatly complicates the regulatory challenge of states. Because of various psychological and social mechanisms that prevent people from recognizing their wrongdoing and encourage them to feel as if they are far more moral, unbiased, and law abiding than they actually are, individuals today are less likely to react, at least not explicitly, to classical legal signals, which they view as directed to other, "bad" people. Similar self-serving mechanisms affecting their perception of social norms and fairness cause people to have very inaccurate views of the normative status of their behavior. Moreover, a great deal of uncertainty surrounds the good-people rationale and, as we will show from the literature, there is clearly more than one type of good person – different people use a variety of different mechanisms to justify their unethical and illegal behavior. We do not yet know how "good" the good

people are in terms of their awareness and ability to control their conduct. Nor can we accurately quantify *ex ante* the ratio of good to bad people in society with regard to any particular behavior. Although we appreciate the need to address the misconduct of good and bad people differently, we do not know the costs of using the "wrong" intervention techniques to deal with various types of bad behavior. Bringing about the needed shifts in regulatory design first requires a shift in the behavioral analysis of law.

1.1 LIMITED COGNITION, LIMITED SELF-INTEREST, AND BEHAVIORAL ETHICS

The past 40 years have seen a dramatic increase in the influence of psychology on the field of economics in general and on the law and economics movement in particular. As a result, significant efforts have been devoted to mapping the flaws in human cognition and examining their implications for how individuals deviate from making optimal decisions.[1] For example, the literature has investigated how irrelevant factors of context, framing, or situation can cause individuals to make decisions that are contrary to their best interest. Daniel Kahneman's book, *Thinking, Fast and Slow*, popularized the concept of two systems of reasoning, which now is at the core of extensive research in behavioral law and economics.[2] Kahneman differentiates between an automatic, intuitive, and mostly unconscious process (System 1) and a controlled and deliberative process (System 2). Although many scholars – for example, Gigerenzer et al.[3] and Kruglanski[4] – have criticized this paradigm, recognition of the role of automaticity in decision making has played an important role in the emergence of behavioral law and economics.

It is essential to clarify at the outset the dramatic difference between the highly popular behavioral law and economics (BLE) and behavioral ethics (BE). BLE is concerned with people's limited ability to make "rational" decisions, whereas behavioral ethics addresses people's inability to fully recognize the ethical, moral and legal aspects of their behavior. How BE and BLE approach self-interest illustrates the main difference between them. BLE assumes that people cannot be fully trusted on their own to make decisions that enhance their self-interest because of the bounded rationality argument – that available information, cognitive ability, and

---

[1] Jolls, C., Sunstein, C. R., & Thaler, R. (1998). A behavioral approach to law and economics. *Stanford Law Review*, 50, 1471–1550. See also Korobkin, R. B., & Ulen, T. S. (2000). Law and behavioral science: Removing the rationality assumption from law and economics. *California Law Review*, 88, 1051–1144; Langevoort, D. C. (1998); Behavioral theories of judgment and decision making in legal scholarship: A literature review. Vanderbilt Law Review, 51, 1499; Jolls, C. (2007). *Behavioral law and economics* (No. w12879). National Bureau of Economic Research.

[2] Kahneman, D. (2011). *Thinking, slow and fast*. New York: Farrar, Straus and Giroux.

[3] Gigerenzer, G., Todd, P. M., & ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

[4] Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review*, 118(1), 97.

time constraints limit individuals' ability to make rational decisions. In contrast, BE focuses on people's inability to recognize the extent to which self-interest in its broader sense affects their behavior. BE assumes that many people's actions are based on self-interest, in that they serve the need to maintain a positive and coherent view of the self. It also accounts for the effect that self-interest has on cognitive processes (e.g., visual perception and memory), as opposed to simply looking at how self-interest affects motivation. Finally, BE is more concerned with how our self-interest affects us implicitly than with how it shapes our explicit choices. In light of these differences, the fact that BLE is so popular within the legal literature[5] while BE is almost entirely ignored[6] is quite counterintuitive.[7]

As I discuss in more detail in Chapter 2 – which focuses on the psychological foundations of behavioral ethics – good people are those who find themselves in situations in which they are not fully aware of the legal, moral, and ethical meanings of their behavior for a combination of reasons.[8] They then engage in motivated reasoning, in which their desires affect the types of information they pay attention to and how they process it.[9] Self-deception also plays an important role in their ability to accurately assess the nature of their actions and motives, causing them to believe they are acting more ethically than they actually are.[10] To use a common

---

[5] For example, Sunstein, C. R. (1999). Behavioral law and economics: A progress report. *American Law and Economics Review*, 1(1/2), 115–157. See also Langevoort, *supra* note 1.

[6] For some comparison of the potential of the two literatures, see Amir, O., & Lobel, O. (2008). Stumble, predict, nudge: How behavioral economics informs law and policy. *Columbia Law Review*, 108(8), 2098–2137.

[7] I discuss this point in Chapter 1, page 6.

[8] For example, see Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, 45(6), 633–644. See also Bersoff, D. M. (1999). Why good people sometimes do bad things: Motivated reasoning and unethical behavior. *Personality and Social Psychology Bulletin*, 25(1), 28–39; Kidder, R. M. (2009). *How good people make tough choices: Resolving the dilemmas of ethical living* (Rev. ed.). New York: Harper Perennial; Pillutla, M. M. (2011). When good people do wrong: Morality, social identity, and ethical behavior, in D. De Cremer, R. van Dijk & J. K. Murnighan (Eds.), *Social psychology and organizations* (p. 353). New York: Routledge; Hollis, J. (2008). *Why good people do bad things: Understanding our darker selves*. New York: Penguin; and Banaji, M. R., & Greenwald, A. G. (2013). *Blindspot: Hidden biases of good people*. New York: Delacorte Press. Many other authors do not use the term but make the same argument in the text (see, e.g., De Cremer, D., van Dick, R., Tenbrunsel, A., Pillutla, M., & Murnighan, J. K. (2011). Understanding ethical behavior and decision making in management: A behavioural business ethics approach. *British Journal of Management*, 22(s1), S1–S4. This is also the view held by Bazerman, M. H., & Tenbrunsel, A. E. (2011). *Blind spots: Why we fail to do what's right and what to do about it*. Princeton: Princeton University Press. This line of scholarship is completely different from the type of research conducted by Zimbardo, P. G. (2007). *The Lucifer effect*. New York: Random House Trade Paperbacks. These works generally try to explain how ordinary people end up doing evil or at least engaging in gross criminal behaviors.

[9] Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480.

[10] Chugh, D., Bazerman, M. H., & Banaji, M. R. (2005). Bounded ethicality as a psychological barrier to recognizing conflicts of interest. In D. A. Moore, D. M. Cain, G. Loewenstein, & M. H. Bazeman (Eds.), *Conflicts of interest: Challenges and solutions in business, law, medicine, and public policy* (pp. 74–95). Cambridge: Cambridge University Press.

example, a mayor will find it difficult admitting to himself that his behavior is driven by anything other than the benefit of the city he runs – even if his specific actions[11] seem to be, on the surface, motivated primarily by his own self-interest.

As discussed in more details in Chapter 2 and especially in Chapter 9 that focuses on implicit corruption, the BE literature has produced many important and counter-intuitive insights with regard to the predictors of unethical behavior. For example, people behave less ethically in groups than when alone[12] and also when they are acting on behalf of other people, rather than for themselves. Another example is that good people might ignore blatant conflicts of interest, having few qualms about accepting tickets to a sports event from a client, although they would shy away from taking a monetary bribe. Individuals who consider themselves to be "good" based on their past behavior may permit themselves to bend the rules (moral licensing) and are more likely to make unethical decisions when time constraints increase.[13] These findings described in the literature pose a substantial challenge to the ability of the state to change the behavior of the public across many domains of law.

As will be developed throughout the book, current research on behavioral ethics could explain a long line of uncooperative behaviors and wrongful conducts, such as breaching contracts due to biased interpretation of the contractual negotiation, engaging in corruption for undermining the effect of self-interest on one's reasoning, employment discrimination due to social cognition processes, and eschewing professional duties of loyalty in various corporate and administrative contexts.

These psychological mechanisms not only amplify the effect of self-interest but also tend to limit people's awareness of the role of self-interest in determining their behavior. Indeed, one of the unresolved issues is the degree to which individuals are aware of their ethical behavior,[14] and BE research has proceeded along several paths that argue different views on this topic. On the one hand, Marquardt and Hoeger showed that individuals make decisions based on implicit rather than explicit attitudes.[15] Along similar lines, when examining the automatic system, Moore and Loewenstein[16] found that the effect of self-interest is automatic, and Epley and

[11] For example, in choosing people he wants to promote, areas in the city he decides to develop, and contractors with whom he interacts. In Chapter 9, the fact that the contribution of the "best interest of the city" is an ambiguous concept is developed.

[12] Weisel, O., & Shalvi, S. (2015). The collaborative roots of corruption. *Proceedings of the National Academy of Sciences*, 112(34), 10651–10656.

[13] Shalvi, S., Eldar, O., & Bereby-Meyer, Y. (2012). Honesty requires time (and lack of justifications). *Psychological Science*, 23(10), 1264–1270.

[14] Hochman, G., Glöckner, A., Fiedler, S., & Ayal, S. (2016). "I can see it in your eyes": Biased processing and increased arousal in dishonest responses. *Journal of Behavioral Decision Making*, 29(2–3), 322–335.

[15] Marquardt, N., & Hoeger, R. (2009). The effect of implicit moral attitudes on managerial decision-making: An implicit social cognition approach. *Journal of Business Ethics*, 85(2), 157–171.

[16] Moore, D. A., & Loewenstein, G. (2004). Self-interest, automaticity, and the psychology of conflict of interest. *Social Justice Research*, 17(2), 189–202.

Caruso[17] concluded that automatic processing leads to egocentric ethical interpretations. However, within BE can be found theories such as Bandura's theory of moral disengagement that maps *post hoc* deliberative and aware self-serving justifications, creating a taxonomy of how people come to explicitly rationalize their unethical behavior.[18]

Another body of literature that stands in contrast to BE is that on limited self-interest, which emphasizes the role of fairness and morality in compliance with the law. A good example is the important line of research that derives from the prosocial account of human behavior (see, e.g., works of Stout[19] and Benkler[20] on prosocial behavior). According to this literature, rational choice models cannot account for our ability to cooperate and engage in prosocial behavior beyond what is in our self-interest.

Both BE and the prosocial behavior literature agree on the need to take a broader view of how self-interest operates relative to traditional economics, and both disagree with the notion that money is the main force motivating people. However, they do not agree on the implications of these assumptions: BE argues that a broad account of self-interest should reveal our tendency toward selfish action, whereas the prosocial literature claims the opposite. In this book, I do not suggest that we look at people's selfish choices to understand their behavior. On the contrary, I offer a more complex view of what it means for a choice to be in one's broader self-interest and how self-interest affects people's understanding of the legal and moral meaning of their behavior.

1.2 THE CONTRIBUTION OF ECONOMICS TO THE DEVELOPMENT OF THE BEHAVIORAL ANALYSIS OF LAW

The contribution of economics to law and psychology, which cannot be overstated, has brought about a shift in focus from the individual to the collective. Before the field of BLE developed, the law and psychology scholarship mostly took a forensic approach, evaluating individuals for the courts, primarily in criminal and family law contexts. Even research exposing biases at work in criminal and civil procedures, which is closely related to research in empirical legal studies (ELS),[21] was often carried out in the context of individuals involved in particular court cases (e.g., jury selection and jury decision making). This orientation has limited the applicability of the traditional law and psychology scholarship to regulatory and legislative contexts.

---

[17] Epley, N., & Caruso, E. M. (2004). Egocentric ethics. *Social Justice Research*, 17(2), 171–187.
[18] Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology*, 71(2), 364.
[19] Stout, L. (2010). *Cultivating conscience: How good laws make good people*. Princeton: Princeton University Press.
[20] Benkler, Y. (2011). *The penguin and the leviathan: How cooperation triumphs over self-interest*. New York: Crown Business.
[21] Rachlinski, J. J., Johnson, S. L., Wistrich, A. J., & Guthrie, C. (2009). Does unconscious racial bias affect trial judges? *Notre Dame Law Review*, 84(3), 1195.

In contrast, BLE scholarship focuses on understanding the behavior of ordinary people in everyday situations, with attention to situational context and the general effect of law on those actions. Many BLE findings have found practical application through communications with regulators, legislatures, and Behavioral Insight Teams (BIT).[22]

BLE incorporates psychological insights into law through an economic lens. At the same time, it ignores many noneconomic areas of psychology, focusing instead on theories related to judgment and decision making. The implications of the limited attention paid to the role of psychological mechanisms in people's behavior are discussed in the next chapter.

This book challenges the excessive focus on cognitive biases at the expense of ethical biases that allow immoral behavior. Whereas the economics literature stresses rationality – that is, the outcome as a utility-maximizing decision relative to preference – I argue that it is the understanding of the importance of non-deliberative decision making that truly matters for legal theory; in addition, it is precisely the nuanced effect of this process on immoral behavior that economics fails to address.

1.2.1 *Demonstration through the "Self-serving Bias"*

The danger of BLE's over-reliance on economics is best demonstrated in the ways its scholarship addresses the self-serving bias. Despite this bias's clear relevance for morality and responsibility and therefore its close relationship to legal theory and enforcement, the BLE literature focuses on it instead as a deviation from rationality. For example, self-serving biases have been held responsible for people's inability to estimate correctly the probability of winning legal battles. Babcock and Loewenstein conducted the most famous study, which showed that self-serving biases operated to reduce the likelihood of people settling out of court.[23] This is a typical BLE finding because it assumes that people make rational decisions – basing their decision to pursue legal action or settle based on their probability of winning. In this case, the self-serving bias suggests a narrow deviation from rationality, causing them to over-estimate their probability of winning. But a much greater problem for the law, one that currently is mostly ignored, is the contribution of the self-serving bias to people's inability to recognize both their own wrongdoing and the dominant role that their self-interest plays in their behavior – which limits their ability to understand why legal action is being brought against them. The law and economics movement has thus limited the richness of the psychology being used in legal scholarship.

---

[22] For a review see Jolls et al., *supra* note 1. See also Korobkin & Ulen, *supra* note 1. See also Halpern, D. (2016). Inside the nudge unit: How small changes can make a big difference. Random House.

[23] Babcock, L., & Loewenstein, G. (1997). Explaining bargaining impasse: The role of self-serving biases. *Journal of Economic Perspectives*, 11(1), 109–126.

The proposed legal perspective is not concerned with whether people are acting rationally. Instead, it is concerned with whether they are at fault, whether their behavior can be modified, and whether something in the situation has affected their ability to recognize their wrongdoing. Understanding these processes of decision making and how they affect questions of motivation, autonomy, and responsibility, rather than how to reach the optimal outcome, should be at the core of the new behavioral analysis of law.

1.3 WHY BEHAVIORAL ETHICS HAS BEEN NEGLECTED IN LAW

As suggested earlier, both the BE and the traditional BLE literatures focus on the automatic processes that underlie people's decision making. However, they have different emphases: BE explores the automaticity of self-interest, whereas BLE examines areas in which automatic decisions undermine self-interest.[24]

Given the importance of intentionality to the law, one would expect behavioral ethics to be more central to legal scholarship than it is today. Yet BE has had less of an impact on the legal arena than has behavioral law and economics. This is primarily because of BE's structural limitations. For example, BE has a relatively large number of founding scholars, whereas BLE has two main ones: Kahneman and Tversky. As a result, BE suffers from the simultaneous development of multiple, competing paradigms, muddling the underlying points on which the literature agrees. These disagreements prevent BE from being able to propose consistent policy recommendations, which is another obstacle to its adoption within the law. Yet another limitation of BE is that it relies to a greater extent than does BLE on dual-reasoning mechanisms, whose concepts of automaticity, awareness, and controllability are difficult to explore and measure. How is it possible to prove that people are unaware or even partly unaware of their selfish intentions? By contrast, classical BLE focuses on suboptimal outcomes, which can be easily examined empirically. This focus places many of the findings of BE at methodologically inferior positions relative to those of BLE.

Finally, another limitation of BE relative to BLE is the greater inability of third parties to recognize the biases of the decision making. When it comes to BLE-related biases such as loss aversion, third parties can more easily recognize the fact that this bias undermines the ability of decision makers to treat loss and profit as similar consequences. By contrast, the main mechanisms in behavioral ethics are related to self-serving biases and motivated reasoning, which contribute to people's reduced ability to recognize their own wrongdoing. Since these mechanisms are self-driven, it is harder for third parties who look at others' bad behaviors to recognize them as

[24] Kish-Gephart, J. J., Harrison, D. A., & Treviño, L. K. (2010). Bad apples, bad cases, and bad barrels: Meta-analytic evidence about sources of unethical decisions at work. *Journal of Applied Psychology*, 95(1), 1. See also Bazerman, M. H., & Gino, F. (2012). Behavioral ethics: Toward a deeper understanding of moral judgment and dishonesty. *Annual Review of Law and Social Science*, 8, 85–104.

"good" people who simply cannot recognize their own wrongdoing. To use a hypothetical example, if a public official promoted a friend, BE suggests a whole array of mechanisms that might bias her ability to recognize the impact of personal familiarity on the objectivity of her decisions.[25] However, for third parties, BE research suggests that they will have trouble believing that the public official did not favor her friends knowingly.[26] Such a gap between the decision maker and third parties also contributes to the reluctance of BLE scholars to adopt BE-based biases as part of the bounded rationality project.[27] Despite the aforementioned limitations, bringing BE into mainstream legal scholarship is both a challenging and rewarding task and it will be the primary occupation of the present book.

1.4 THE GIST OF THE BOOK

As alluded to in the previous paragraphs, in this book, I aim to create a new branch of scholarship that focuses on the rule of law in a world populated by individuals with different levels of awareness of their own unethicality. This book is based on the assumption that many of the current directions in legal enforcement research, especially with regard to 'ordinary unethicality' miss important elements of both behavioral and legal methods and theories. It challenges the ability of states to systematically account for non-deliberative, unethical human behavior given a legal system based largely on either sanctions or moral messages, both of which assume some level of calculation and deliberation. The legal literature on enforcement needs to undergo a major revision in its approach to the regulation of intellectual property, employment discrimination, conflict of interest, and many other legally relevant behaviors that people engage in for multiple reasons and with limited awareness of their full legal and moral meaning. In such contexts, the BE approach is especially potent and needs to be taken into account. This change in perception creates many new challenges from a regulation and enforcement perspective, as it is unclear to what extent current legal instruments could be seen as effective in curbing misconducts conducted by people limited awareness to the full meaning of their own behavior. The focus of the book is to explore the ability to create regulatory and enforcement tools that will be able to target people who differ in their self-awareness to wrongdoing.

As suggested, the book criticizes the behavioral-legal scholarship for overemphasizing rationality and cognitive biases at the expense of non-deliberative choice and

---

[25] See discussion in Chapter 2 on the objectivity bias.

[26] See discussion in Chapter 9 on implicit corruption.

[27] Compare with the argument made in Soltes, E. (2016). *Why they do it: Inside the mind of the white-collar criminal*. New York: PublicAffairs, where convicted white color criminals report they were unaware at the time that their behavior was unethical or illegal. The vast majority of people find it very hard to believe that those people indeed did not know what they were doing.

ethical biases. However, as is shown throughout the book, the move to dual-reasoning theories should not lead to a categorical rejection of deterrence and morality. In fact, the reverse is true: one of the arguments developed in later chapters is that traditional enforcement mechanisms have more than one type of effect on people; therefore, the current fascination with "nudges"[28] as a means of changing behavior, along with the abandonment of traditional intervention mechanisms, is misguided.

In latter chapters, I examine the new insights derived from behavioral ethics, a relatively overlooked area in current legal research, which help identify many mechanisms that prevent people from fully recognizing the wrongfulness of their behavior. At a conceptual level, the book revises some jurisprudential concepts related to choice, responsibility, and autonomy in light of growing knowledge about the role of non-deliberative choice in human behavior. Based on these insights, I revisit many of the existing behavioral paradigms of legal regulation and enforcement and conclude by presenting a multidimensional taxonomy of legal doctrines and of the various instruments that states can use to modify human behavior. I recommend certain changes that legal scholarship on enforcement needs to make to remain relevant in the face of recent behavioral research and regulatory changes.

Such a change in focus would greatly affect the design and enforcement of laws and regulations in many legal domains. For example, how can we justify the use of deterrence in light of the "blind spot" argument (i.e., ethical unawareness) advanced by scholars such as Bazerman and Tenbrunsel as well as Banaji and Greenwald?[29] How can we understand the legal responsibility of organizations given what we know about situational cues of unethicality? How should we think of nudges when our goal is to increase ethicality, rather than improving the available choices, although only the latter are in the long-term interest of individuals? How are we to understand the *Why People Obey the Law* project of Tom Tyler,[30] which is based on self-report and explicit accounts of fairness, in light of the writings on moral intuition by Haidt[31] and on moral identity by Aquino?[32] Should we ascribe a new meaning to legal ambiguity, given its contribution to such processes as the moral wiggle room and self-deception? Can

---

[28]   The concept of nudges, which is discussed in Chapter 4, was advanced in the 2008 book by Richard H. Thaler and Cass R. Sunstein (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New York: Penguin, 2008. A nudge is a simple intervention, such as changing the default setting in decision making, that policy makers can institute to change people's behavior with a limited need for them to make any deliberative choice.

[29]   See Sunstein, *supra* note 5. See also Bazerman & Tenbrunsel, *supra* note 8; Banaji, M. R., & Greenwald, A. G. (2016). *Blindspot: Hidden biases of good people*. New York: Bantam Books.

[30]   Tyler, T. R. (2006). *Why people obey the law*. Princeton: Princeton University Press.

[31]   Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814.

[32]   Aquino, K., & Reed II, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83(6), 1423.

states use enforcement mechanisms that distinguish between intentional and situational wrongdoers?

In general, I argue that we should separate situations of specific individuals – where we need to define *ex post* the level of responsibility of a given individual who is on trial given his or her own limited awareness – from situations where we examine *ex ante* how to mobilize a given population, where our focus is on the collective. The first type of situation is the traditional view of law, but the fact that current studies show that ethical awareness is limited might not be enough to lead to a normative change without more research. However, when it comes to *ex ante* intervention, even when we cannot fully determine the strength of the non-deliberative component in people's ethical motivation, we are able to predict that this component is likely to change the behavior of an unknown proportion of the population and hence should affect the *ex ante* design of law.

In subsequent chapters, I attempt to bridge the gap between the new findings of the behavioral ethics approach to behavior and existing methods used to modify behavior. The new behavioral approaches to law enforcement assume that individuals are motivated to engage in illegal conduct by more than the pursuit of material self-interest. These approaches collide with the traditional outlook, requiring a broad theoretical and empirical comparison of both traditional enforcement mechanisms and nontraditional measures to understand how states may be able to cope with bad deeds carried out by people with a variety of motivations and levels of awareness. I explore the meaning of these variations across people, types of behavior, and legal doctrines.

This book explores the pros and cons of each regulatory tool available to government using an instrument-choice perspective based on the extensive knowledge we already have on the behavioral implications of each tool. This analysis assesses the advantages of both traditional and nontraditional approaches to legal enforcement in addressing both general enforcement dilemmas and contexts of fighting corruption and discrimination.

1.4.1 *The Challenge to Legal Enforcement Posed by Behavioral Ethics*

The underlying assumption of BE regarding the complex role played by the "self" in ethical decision making is clearly problematic for legal theory. BE proposes that many of the claims about the responsibility of individuals as moral agents for their actions neglect the impact of the situation in which the decision-making process is taking place. It may be that the main driver of the individual's behavior is the situation and not the individual's current self-view. Furthermore, the automaticity of the self-enhancement process creates a "responsibility gap" for the individual who is not completely aware of the ethicality of his or her actions and therefore cannot be held responsible for them. A possible way of bridging this gap is through nudges and

by designing the situation so that it enhances moral awareness and calculated decision making.

The argument that I develop throughout the book is that the current level of knowledge that BE is able to provide is limited, especially with regard to important questions from a legal perspective; therefore, it is not able to provide policy makers with a complete list of recommendations on how laws should be changed. We lack sufficient knowledge about individuals' awareness of the unethicality of their behavior and their ability to control these ethical biases.[33] The psychological and social mechanisms, which I describe in detail in the next chapter, paint a complex picture of human character according to which people mostly seek to promote their self-interest as long as they can feel good about themselves.[34] Based on this paradigm, if we allow people to choose how to behave, many good people might resort to self-deception mechanisms, such as moral disengagement or elastic justification, and take advantage of others' trust to shirk their responsibilities, engage in dishonest behavior, or violate the law.

1.4.2 *Toward a Broader Perspective of the Regulation of Good People*

One of the most difficult challenges this book addresses is how interventions can increase ethical behaviors if most unethical actions are done unconsciously. For example, Bazerman and Banaji, two of the leading scholars of ethical decision making, argue that incentives and similar concepts fail to correct a large portion of unethical behaviors because "such measures simply bypass the vast majority of unethical behaviors that occur without full conscious awareness of the actors, who engage in them."[35] If we accept this argument, we can challenge enforcement methods that focus on external measures and incentives to control unethical behavior because they ascribe an unjustifiably key role to self-control, autonomy, and responsibility for action. One of the main shortcomings of the "good-people" literature is the gap between what we know about the dominant role of System 1 in ethical decision making and about what policy makers can do to curb thoughtless and unethical behaviors. Evidence of the automaticity of unethicality suggests that a new approach is required to create effective enforcement methods across all fields of legal regulation.

This recognition of the need for a new approach to enforcement lies at the heart of this book. I argue here that the state needs to differentiate enforcement methods: targeting "traditional" misconduct with traditional measures and nontraditional, only partially aware, misconduct with different types of interventions. Throughout

---

[33] For conflicting results with regard to physiological indications of dishonesty among people, see Hochman et al., *supra* note 14.

[34] See Mazar et al., *supra* note 8.

[35] Bazerman, M. H., & Banaji, M. R. (2004). The social psychology of ordinary ethical failures. *Social Justice Research*, 17(2), 111–115.

this book, and particularly in Chapter 2, I describe mechanisms to address the nontraditional misconduct of people.

The book examines states' and organizations' ability to prevent people from engaging in uncooperative behavior, such as wrongful conduct, breach of contracts, and eschewing of professional duties, through traditional methods and compares those methods' effectiveness and limitations to behaviorally informed enforcement mechanisms such as the nudge approach, framing, expressive law, and procedural justice. After discussing the pros and cons of the various intervention mechanisms, the book recommends practical steps for legal policy makers to optimize their regulatory and enforcement efforts to influence both the deliberative and non-deliberative components of behavior.

These practical recommendations are based on a coherent account of the person that the law tries to affect and control, as well as an integrated consideration of many unresolved theoretical questions, such as the following: How much can we know *ex ante* about the awareness, controllability, and modification of the behaviors of good people? How can we know that their goodness is genuine and not fake? Are morality and traditional enforcement practices, such as deterrence, effective in curbing behaviors that are only partly deliberative? Can states regulate good and bad people by using different enforcement methods? Should the nudge approach, which avoids direct communication between the state and the people it regulates, replace all other intervention methods? Do we know to what degree the sustainability of behavioral change and autonomy is reduced when we abandon traditional inter-vention methods?

Some of these questions have been the subject of empirical and theoretical studies, and the book draws on extensive empirical research that others and I have conducted. For example, I have studied empirically the effect of social norms on the perception of legality in the context of intellectual property, the effect of incentives on people's intrinsic and extrinsic motivations in the area of environmental protection and whistle-blowing laws, and the effect of legal uncertainty on the compliance and performance of people with different motivational back-grounds. The book moves one step further and fills the gaps unresolved by these and other studies by answering this key question: How much does the behavioral analysis of law (which studies deterrence, legitimacy, procedural justice, and the expressive function of the law) have to offer to improving legal compliance, as understood today, given what we know about the role of automaticity in legal compliance? Although earlier studies have contributed to the body of knowledge of the behavioral analysis of law, they fall short of offering a coherent behavioral and normative picture of the person we are trying to regulate and of answering the questions raised earlier.

1.4.3 *Behavioral Ethics and the Instrument-choice Literature*

An additional goal of the book is to integrate the growing interest in BIT with the increasing use of nontraditional measures, such as nudges, thereby addressing the debate in the literature over "legal instrument choice"[36] and experimental legislation.[37] The change in regulatory instruments follows directly from the recognition of people's bounded rationality. One response to the greater appreciation of the role of non-deliberation in decision making is a move from a command-and-control approach to softer types of regulation.[38] Traditional enforcement mechanisms used by states worldwide are based on the assumptions that people actively chose to engage in "bad" behaviors and that techniques such as incentives can be used to change those choices.

However, recent research shows that much of that behavior is engaged in unconsciously. This book proposes the most effective ways to change behavior by accounting for the effect of both traditional and nontraditional enforcement methods on public trust, legitimacy, and the perceived rule of law. In the context of instrument choice, it addresses the following questions: How do the recommendations of BITs affect people with different modes of reasoning and with different motivations regarding the law? How do modes of reasoning interact with previously shown effects of motivations on legal compliance? What are the long-term effects of BITs proposals on people's perception of responsibility and autonomy? Can we find connections between knowledge of behavioral ethics on the part of the law and legal concepts such as negligence, acting knowingly, and intentionality? Can we identify connections between people's motivations regarding the law and the likelihood of engaging in ethical biases? Can we find the optimal balance between traditional methods and nontraditional ones? Can we be sure that, when looking at the most effective legal intervention for a given situation, we measure not only short-term effects but also factors such as legitimacy, perception of the rule of law, and durability of behavioral changes?

For the interaction between behavior-based regulation and the broader concept of law to be meaningful, it is necessary to identify the steps that would allow psychological knowledge to be generalized to the societal level rather than remaining at the individual level. The ideal behavioral approach to law, advocated by this book, must be sensitive to various normative and institutional factors such as trust, legitimacy, and legal culture. Only by combining the behavioral approach with institutional and normative ones can we create a coherent theoretical framework for nontraditional instruments that states can use to achieve greater success than with earlier,

[36] See, for example, Bemelmans-Videc, M. L., Rist, R. C., & Vedung, E. O. (Eds.). (2011). *Carrots, sticks, and sermons: Policy instruments and their evaluation* (Vol. 1). New Brunswick, NJ: Transaction Publishers.

[37] Gubler, Z. J. (2014). Experimental rules. *BCL Review, 55,* 129; Gersen, J. E. (2007). Temporary legislation. *University of Chicago Law Review, 74*(1), 247–298.

[38] Lobel, O. (2004). The renew deal: The fall of regulation and the rise of governance in contemporary legal thought. *Minnesota Law Review, 89,* 342.

narrower approaches. Integrative behavioral research, which explores and analyzes the approaches that government should follow to regulate various types of unethicality in society, can provide policy makers with the methods they need across all legal contexts, going beyond the current focus on energy savings, pension planning, and food consumption; these methods can reach into areas where traditional enforcement methods have failed to produce sustainable change because of their limited focus.[39]

The "Law of Good People" paradigm requires revisiting many of the existing behavioral models of legal regulation and enforcement, which for the most part have relied on the assumption of deliberateness and rationality. These questions need to be answered: What is the optimal use of incentives? Should we replace traditional enforcement mechanisms with nudge interventions? What should be our attitude toward the expressive function of the law, the effect of fairness, or the interaction between incentives and fairness? Furthermore, in contrast to behavioral economics, which deals with biases that prevent people from behaving in a desirable way (e.g., saving more, eating more healthfully), from the BE perspective, many people behave in a way that they consider to be desirable, even after they have had time to reflect on their behavior; others, however, do not.

Yet the need to regulate both good and bad people long preceded the BE revolution. States deal with a world in which people have different motivations to comply with the law, mostly because of their differing levels of internalized moral and legal norms. The solution to dealing with a variety of people lies in the common denominator approach, along with a nuanced use of incentives to prevent crowding-out effects. But when it comes to differences between people's level of deliberation, it is not clear that there is a common denominator at work. And even with the spotlight aimed at the new approach, the previous dichotomies of extrinsic and intrinsic compliance motivations remain relevant, maybe even more so than before. We are now facing the need to regulate people across two dichotomies, which are not necessarily orthogonal: their internalization of norms and their mode of reasoning.

Despite our growing understanding of "good people", no one-size-fits-all policy suggests itself. For the legal policy maker to be able to use the rich knowledge about people's bounded ethicality, we need to create a multidimensional taxonomy of legal doctrines and of the various instruments that states can use in their attempt to modify human behavior. The deviation from the assumption that an actor did wrong because he or she had planned on doing so is justified only in some legal doctrines and only with regard to certain situations. Being able to recognize *ex ante* the areas in which people's lack of moral awareness is expected to be significant can change the balance of the tools that should be used.

---

[39] Bubb, R., & Pildes, R. H. (2014). How behavioral economics trims its sails and why. *Harvard Law Review*, 127(6), 1637.

1.4.4 *Limitations in the Current Behavioral Ethics Literature*

Several limitations in the behavioral ethics literature prevent it from fully incorporating and applying the research on non-deliberative choice and then using the results to improve the ethicality of society in many important domains of life. There is almost no discussion in BE of such concepts as the controllability of non-deliberative choice and awareness of its effect on behavior. The absence of feedback from the applied behavioral sciences limits the ability of basic theoretical science to provide clear answers regarding aspects of non-deliberative choice.[40] In addition, little comparative research has been conducted on the efficacy of various intervention methods in various contexts. The absence of research leads to a lack of serious attempts in the legal and behavioral literature to understand the mechanisms on which intervention methods are based. Behavioral research does not account for trade-offs between the methods and therefore provides no normative guidelines for policy makers.[41] Finally, because of an absence of substantial interaction with legal scholarship, the fields of behavioral engineering and mechanism design treat law in a simplistic way, ignoring the normative complexities and goals embedded in each legal doctrine and painting all legal doctrines with a broad brush.[42]

There is a great need for a richer view of the interaction between law and human behavior that accounts for the effect of legal intervention on good and bad people alike. There also needs to be a deeper understanding of trade-offs, which should be taken into account when evaluating the effectiveness of government intervention in changing behavior. An example of such context sensitivity appears in this book when I address the area of ethical decision making. Empirical results show that accountability is effective in undermining unconscious biases.[43] However, this might not be the case for bad people who are looking for ways to rationalize their intentional bad behaviors. The inability to predict the effect of various legal interventions on behavior demonstrates the need for evidence-based behavioral-legal scholarship. Legal scholarship must recognize that behavioral findings are not merely on the sidelines but are at the heart of the theory and practice of legal enforcement. It must also demonstrate to scholars in the behavioral and public

---

[40]   The assumption of a one-way influence is problematic on many grounds. In the area of non-deliberative choice, the book suggests that understanding what intervention methods work in what circumstances can help basic science gain clarity regarding the interplay between deliberative and non-deliberative choice.

[41]   Feldman, Y., & Lobel, O. (2015). Behavioural trade-offs. In A. Alemanno & A.-L. Sibony (Eds.), *Beyond the land of nudges spans the world of law and psychology in nudge and the law*. Oxford: Hart.

[42]   For example, in many economics papers about contracts or employment discrimination, the legal doctrine is presented rather naively. The nudge approach often ignores alternative solutions offered by the doctrine itself.

[43]   Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, 125(2), 255.

policy fields that law is a unique area that cannot be overlooked.[44] To regulate behavior in a comprehensive way, legal scholarship must adopt an integrative methodological and theoretical approach to the deliberative and non-deliberative predictors of behavior.

Based on the new insights generated by BE, current research seeks to demonstrate that the methods of behavioral-legal scholarship are no longer sufficient. For example, the main studies in the "why people obey the law", mentioned above, assumes for the most part a deliberative process.[45] Likewise, experiments in which participants are required to play cooperation games assume that people readily recognize their self-interest and the public interest. At the same time, the literature on non-deliberative choice ignores the possible effect of compliance on factors such as perception of legitimacy and public trust, as well as cultural and institutional constraints. Furthermore, lab research lacks the required methodological focus of field experiments that provide external validity, which is so much more important for law than for psychology.

For example, one of the main techniques that "good people" can use to self-justify unethical behavior is to engage in a biased interpretation of the legal requirements they must follow. Research on corruption and conflict of interest has studied numerous examples of situations in which people who exhibit professional and moral responsibility have allowed their self-interest, possibly without full awareness, to prevail over fulfilling their duties.[46] The existing literature on contractual performance decisions and framing focuses on the dichotomous choice: to breach or not to breach. However, in a study coauthored with Teichman and Shur,[47] I argued that the focus should not be on whether people choose to comply with contractual obligations, but on their decision to interpret the contract in a self-serving way. This is in contrast to the work of Wilkinson-Ryan and Baron; describing to their participants the promisor's decision to breach a contract, they simply stated, "He decides to break his contract in order to take other, more profitable work."[48] Such studies implicitly assume that choices are made in reference to clear contractual

---

[44] Dagan, H., Kreitner, R., & Kricheli-Katz, T. (forthcoming). Legal theory for legal empiricists. *Law and Social Inquiry*.

[45] See discussion in Chapters 3, 4, and 5 on the expressive effect of law – where the language of the law changes the social meaning of behavior.

[46] There is a wealth of research on the prevalence of conflicts of interest in almost every field. See, for example, Rodwin, M. A. (1993). *Medicine, money, and morals: Physicians' conflicts of interest.* New York: Oxford University Press. See also Thompson, D. F. (1993). Understanding financial conflicts of interest. *New England Journal of Medicine*, 329, 573–576.

[47] Feldman, Y., Schurr, A., & Teichman, D. (2013). Reference points and contractual choices: An experimental examination. *Journal of Empirical Legal Studies*, 10(3), 512–541.

[48] Wilkinson-Ryan, T., & Baron, J. (2009). Moral judgment and moral heuristics in breach of contract. *Journal of Empirical Legal Studies*, 6(2), 405–423 at 413. See also Wilkinson-Ryan, T., & Hoffman, D. A. (2010). Breach is for suckers. *Vanderbilt Law Review*, 63, 1003, at 1029 (using precisely the same phrase to describe the decision to breach).

obligations. This book instead focuses on the arguably more common situation of how to interpret an ambiguous obligation.

Similarly, much of the current literature on morality and legal compliance examines people's moral judgment but ignores the role of moral intuition and the fact that people might engage in motivated reasoning. The classical approach assumes that people consider the situation, recognize the moral conflict, and then decide what to do. This approach ignores that fact that they decide what seems to be the right thing to do based on their highly motivated perception of the situation. Their behavior may be immoral, but they still view themselves as moral people because they frame the situation in such a way that it "allows" immoral behavior. Clearly, people's self-image of being cooperative or moral is not based on acting morally. Many studies in the social cognitive literature discuss mechanisms that either prevent people from knowing in advance that they are violating the law or enable them to develop an *ex post* approach that uses various strategies to change people's perception of the wrongfulness of their behavior. Ariely et al. have shown that people do not believe that it is legitimate to cheat more if one is financially deprived.[49] But when they were manipulated into thinking that they were deprived (by receiving a smaller amount of money in a game), they were quick to start cheating. This finding shows the importance of explicit judgment not only relative to implicit judgment but also to actual behavior. The methodological observation of Greenvald and Banaji on the power of implicit judgment may have even stronger force: because people love themselves so much, there is no reason for them to admit to themselves that they behave amorally.[50]

1.5 BRIEF OUTLINE OF THE CHAPTERS

Chapter 2: Behavioral Ethics and the Meaning of Good People for Legal Enforcement

In this chapter, I review the work of scholars such as Bazerman, Banaji, Ariely, Gino, Haidt, Bereby-Meyer, Shalvi, Rand and others on deliberative and non-deliberative mechanisms that people use to promote their self-interest and that result in good people doing bad things. I discuss the relevance to BE of several theoretical mechanisms: moral disengagement, embodiment, self-deception, moral licensing, automaticity of self-interest, moral hypocrisy, elastic justification, ethical fading, and the dishonesty of honest people. I also describe various

---

[49] Sharma, E., Mazar, N., Alter, A. L., & Ariely, D. (2014). Financial deprivation selectively shifts moral standards and compromises moral decisions. *Organizational Behavior and Human Decision Processes*, 123(2), 90–100.

[50] Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1), 4.

phenomena that challenge the current regulatory approach followed by most states.[51]

In the last part of this chapter, I offer a few words of caution to states that wish to use the knowledge of behavioral ethics to modify legal policy making. The field is relatively young, and it is not yet able to answer important questions with regard to ethical biases, such as their internal mechanisms, the awareness of the existence of biases, and the variation between people in their ability to overcome those biases. However, I strongly believe that legal scholarship cannot wait for a consensus to be reached on those questions. Although we do not know the percentage of good versus bad people, the extent of involvement of the automatic system in decision making, or the level of awareness of its role, the recognition that a substantial portion of the population engages in non-deliberative choices is enough to shift the normative debate. The chapter concludes that knowing the ratio of people of each type and the exact level of people's awareness is secondary in importance to the fact that such variation exists and that legal policy making has to recognize the need to adopt more than one type of intervention to deal with different modes of awareness. We cannot afford to wait until we know more before we act.

Chapter 3: Resisting Traditional Enforcement Mechanisms

In Chapter 3, I reexamine some behavioral theories that explain people's motivation to follow the law. One of the key arguments of this book is that to deal with good people, we do not need to abandon everything we know about legal enforcement; we just need to revise our current understanding of how traditional intervention methods can influence different types of people. Chapter 3 focuses on traditional enforcement methods that were not intended to take into account dual reasoning and the need to deal with non-deliberative choices. Nevertheless, many of those traditional paradigms are highly sensitive to the behavioral revolution. Among the concepts reviewed in this chapter are deterrence, morality and fairness, incentives, social norms, and the expressive function of the law.

This chapter examines the techniques that governments use to regulate behavior.[52] Traditional intervention methods discussed in the literature include incentives (all forms of penalties, fines, rewards,[53] and other external

---

[51]   In the next chapter, I examine the relevancy of paradigms such as egotism (sharing first names, birthdays); embodiment (washing hands, closing one eye, carrying weight); food consumption (drinking coffee, glucose studies); priming the ten commandments, eyes, faces, money, dirty money, and various situational cues that trigger compliance (teddy bears).

[52]   There is a huge literature on the typology of regulatory approaches that I do not discuss here. I review many of them in my work with Lobel, see Lobel & Feldman, *supra* note 41.

[53]   For the literature on this topic, see Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2), 169–217. See also Feldman, Y., & Lobel, O. (2009). Incentives matrix: The comparative effectiveness of rewards, liabilities, duties, and protections for reporting illegality. *Texas Law Review*, 88, 1151. For a comparison of the efficacy of incentives that use

measures[54]) and more intrinsic measures such as fairness-, legitimacy-, and morality-based interventions;[55] social norm-based interventions;[56] and the expressive function of the law, which shapes the social meaning of behaviors.[57] Another technique reviewed in this chapter focuses on disclosure and transparency, which assume that people engage in deliberation, so that given enough information, they will make the right decision or will avoid making the wrong one if what they do is open for everyone to see.

The theoretical and critical part of this chapter focuses on examining how formal enforcement methods should be modified, given that people are not fully deliberative in their decisions to disobey the law or breach their contracts. I show how various subliteratures in the behavioral approach to law (e.g., social norms, compliance motivation, the perceived role of self-interest, the non-instrumental effects of law) should be revised in light of the new knowledge regarding dual-system reasoning and BE. The chapter concludes with a description of the new challenges that policy makers face as a result of recent findings about ethical decision making (e.g., do people who do not want to discriminate react to penalties?) and the development of various behaviorally informed approaches to legal compliance.

The traditional line of research on the role of fairness and morality in legal research could be traced to reasoning advocated by scholars such as Kohlberg[58] on moral development or Kelman[59] on compliance. Kohlberg's ideas, which were incorporated into legal theory, maintain that people clearly recognize that they are facing a moral dilemma, and the question is only what kind of moral rule they use in any given context. By contrast, I argue that people do not regard much of the bad behavior in which they engage as bad behavior at all. Various processes, operating on various levels of self-awareness, help them interpret their behavior as being

---

different labels (e.g. fines, rewards), See Feldman, Y., & Perez, O. (2012). Motivating environmental action in a pluralistic regulatory environment: An experimental study of framing, crowding out, and institutional effects in the context of recycling policies. *Law & Society Review*, 46(2), 405–442 for a comparison of the effects of sanctions versus taxes in environmental contexts.

[54] Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. *The Journal of Economic Perspectives*, 25(4), 191–209.

[55] See Tyler, *supra* note 30. For a review, see Feldman, Y. (2011). Five models of regulatory compliance motivation: Empirical findings and normative implications. In D. Levi-Faur (Ed.) (2013). *Handbook on the politics of regulation* (pp. 335–347). Cheltenham: Edward Elgar.

[56] McAdams, R. H. (2000). Law and Economics Working Papers Series. *Oregon Law Review*, 79, 339–390. See also Feldman, Y., & Nadler, J. (2006). The law and norms of file sharing. *San Diego Law Review*, 43, 577.

[57] Sunstein, C. R. (1996). On the expressive function of law. *University of Pennsylvania Law Review*, 144(5), 2021–2053. See also Feldman, Y. (2009). The expressive function of trade secret law: Legality, cost, intrinsic motivation, and consensus. *Journal of Empirical Legal Studies*, 6(1), 177–212.

[58] Kohlberg, L. (1981). *The philosophy of moral development: Moral stages and the idea of justice*. New York: Harper & Row.

[59] Kelman, H. C. (1958). Compliance, identification, and internalization: Three processes of attitude change. *Journal of Conflict Resolution*, 2(1), 51–60.

either legal and ethical or justifiably illegal and unethical. Some mechanisms even prevent people from recognizing that they are facing a moral dilemma.

The BE perspective also differs from that of most of the compliance motivation literature, including Kelman's basic paradigm on compliance and Tyler's procedural fairness approach. According to Tyler's approach, the main difference is between people with an extrinsic versus those with an intrinsic commitment to obey the law. I claim instead that in many cases, people do not make an informed decision about the right way to act based on their level of commitment to the law. The approach I advocate suggests that the effect of morality on law operates in a completely different way, at least in some cases of legal noncompliance. The core argument of BE is that in a considerable number of cases, people do not engage in any form of deliberative moral reasoning before deciding whether or not to obey the law. Many bad deeds are not seen as such by the people who commit them; therefore, they are not always aware that they might be about to perpetrate a moral wrong. It follows that the focus should be on identifying both the situational and the personality characteristics that will increase the likelihood that people will recognize the moral flaw in their behavior.[60]

According to the traditional approach, certain techniques, most notably incentives,[61] can be used to change the behavior of people so that they do not act in an unethical manner in situations relevant to legal and public policy. In the past three decades, this approach, based on the neoclassical economic doctrine of rational choice, has been challenged by theories based on the behavioral approach to human judgment and decision making. Various alternatives and modifications in regulating human behavior, going beyond simple incentives, have been offered over the years, including some in my own research. For example, the following interventions have been proposed: those that change the wording of incentives (to make people more likely to consider their ethical behavior), that increase legitimacy (by mandating employee voice procedures such as pay talks and hearing prior to termination, in the workplace), that account for crowding out (e.g., when people would rather do things without being compensated for them), that increase the sensitivity to cognitive limitations (e.g., in examining how people engage in aggressive interpretation when contractual obligations are framed as a potential loss rather than a potential gain).[62] Nonetheless, the challenge to legal

---

[60]   In this vein, a work conducted on the scale of rule conditionality is Fine, A., van Rooij, B., Feldman, Y., Shalvi, S., Scheper, E., Leib, M., & Cauffman, E. (2016). Rule orientation and behavior: Development and validation of a scale measuring individual acceptance of rule violation. *Psychology, Public Policy, and Law*, 22(3), 314.

[61]   Camerer, C. F., Hogarth, R. M., Budescu, D. V., & Eckel, C. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. In B. Fischhoff & C. F. Manski (Eds.), *Elicitation of preferences* (pp. 7–48). Dordrecht: Springer.

[62]   Feldman, Y. (2011). For love or money? Defining relationships in law and life: The complexity of disentangling intrinsic and extrinsic compliance motivations: Theoretical and empirical insights from the behavioral analysis of law. *Washington University Journal of Law and Policy*, 35(1), 11–51.

enforcement posed by the BE perspective has not been explored in this context, mainly because of the dominance of economics in the interplay between law and psychology.

Chapter 4: Revisiting Non-formal Enforcement Mechanisms

This chapter examines new nontraditional, soft enforcement methods that take into account people's limited awareness and cognition; these methods include the nudge (an intervention that changes behavior without creating economic incentives or banning other possibilities),[63] de-biasing (a group of doctrines and methods used to overcome biased thinking),[64] accountability, and reflection (which requires individuals to explain why they made a certain decision after making it).[65] I also emphasize the effectiveness of these new behavioral measures, which were developed to deal with cognitive biases, in overcoming ethical biases. This chapter concludes with a description of experimental work comparing the efficacy of explicit and implicit types of interventions on how people behave in subtle conflict of interest situations.

The chapter reviews the studies that have demonstrated the limitations of each intervention when dealing with people who lack a full awareness of their behavior; these limitations are evident in the few famous failures of the nudge-based approach. For example, initially the default rule – having to opt in versus out when donating – seemed to be an effective nudge in increasing rates of organ donation; however, later studies showed that the nudge was actually much less effective.[66] The message, "Save more tomorrow," shown to be a strong nudge in pension savings, also turned out to be less effective in the long term.[67] Generally speaking, giving people full information proved to be problematic;[68] de-biasing was found to produce limited results,[69] as was disclosure of conflicts of interest, which in many contexts ended up having the opposite effect from the desired one.[70] Masking personal information in

---

[63] See Jolls et al., *supra* note 1.

[64] Jolls, C., & Sunstein, C. R. (2006). The law of implicit bias. *California Law Review*, 94(4), 969–996.

[65] See Lerner & Tetlock, *supra* note 43. There are many other techniques that have been studied in recent years, such as signing first and teddy bears, reviewed and assessed in Feldman, Y. (2014). Behavioral ethics meets behavioral law and economics. In E. Zamir & D. Teichman (Eds.), *Oxford handbook of behavioral law and economics* (pp. 213–241). New York: Oxford University Press.

[66] Orentlicher, D. (2008). Presumed consent to organ donation: Its rise and fall in the United States. *Rutgers Law Review*, 61, 295.

[67] See Bubb & Pildes, *supra* note 39, at 1637.

[68] Ben-Shahar, O., & Schneider, C. E. (2011). The failure of mandated disclosure. *University of Pennsylvania Law Review*, 159(3), 647–749.

[69] Larrick, R. P. (2004). Debiasing. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 316–338). Oxford: Wiley-Blackwell.

[70] Cain, D. M., Loewenstein, G., & Moore, D. A. (2005). The dirt on coming clean: Perverse effects of disclosing conflicts of interest. *Journal of Legal Studies*, 34(1), 1–25.

hiring applications was found to be more effective in reducing biases against minorities than against women.[71] Thus predicting when and how to change behavior through non-traditional mechanisms such as nudges has proven to be more difficult than previously assumed.

Chapter 5: Social Norms and Compliance

Chapter 5 focuses on the social norms in legal compliance. This chapter was separated from Chapters 3 and 4 due to its interaction with both traditional and nontraditional means of intervention. Hence, this relatively short chapter outlines the main literature on the role of social norms, with particular focus on non-deliberative processes in how people perceive social norms and in how that affects their behavior.

Chapter 6: Are All People Equally Good?

In this chapter, I examine several factors that might explain the variations among people in their level of implicit and explicit legally relevant behaviors. Because personality scales are an important measure of such variations, I review several scales, including those relating to moral identity,[72] level of moral disengagement,[73] and moral firmness,[74] as well as context-specific measures, such as racism, which are based on the implicit association test.

Second, following a critical assessment of the relevance of data accumulated in personality research to legal compliance, given the great variation in the factors that could be correlated with that compliance, I examine the research in economics, psychology, and law concerning the differences between the intrinsic and extrinsic motivation of individuals; this is another way to differentiate among people.[75] Variations in compliance usually depend on the particular content of the law; some people may be highly motivated to obey certain laws but not have the same level of intrinsic motivation to obey other types of laws.[76]

---

[71]   Krause, A., Rinne, U., & Zimmermann, K. F. (2012). Anonymous job applications of fresh Ph.D. economists. *Economics Letters*, 117(2), 441–444.

[72]   Reed II, A., & Aquino, K. F. (2003). Moral identity and the expanding circle of moral regard toward out-groups. *Journal of Personality and Social Psychology*, 84(6), 1270.

[73]   Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, 3(3), 193–209.

[74]   Shalvi, S., & Leiser, D. (2013). Moral firmness. *Journal of Economic Behavior & Organization*, 93, 400–407.

[75]   For the main economic model, see Benabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *Review of Economic Studies*, 70(3), 489–520. For a review of implications to legal theory, see Feldman, *supra* note 62.

[76]   Although, as explained in this chapter, there are some general tendencies to obey the law. These tendencies are supported also with regard to the citizenship approach to legal compliance whereby people obey the law simply because it is the law. It is interesting to examine whether people who obey laws simply because they are laws should be seen as intrinsically committed individuals.

To understand the "law of good people" paradigm, it is necessary to examine whether there is a correlation between people's level of motivation and their personality traits. For that reason, I focus on more specific personality traits that influence people's general tendency to obey or disobey the law. I discuss the work of Glockner et al.,[77] who compared situational and individual factors related to legal compliance with a focus on self-control. In addition, I review a new scale developed by me in a joint work with Fine, Van Rooji,[78] and others on people's ability to find excuses for violating the law.

Chapter 7: Pluralistic Account of the Law: The Multiple Effects of Law on Behavior

The pluralistic account of the effect of the law on behavior is based on several assumptions developed in this chapter. Most people obey the law for multiple reasons, but people are more likely to experience and report their more noble motivations. Dual reasoning greatly reduces our ability to measure the "true" effect of the law. The law, according to this view is multi-faced and must communicate with different populations at the same time. Some aspects of the law serve multiple functions simultaneously, as for example, the expressive effects of punishments.

Chapter 8: Enforcement Dilemmas and Behavioral Trade-offs

Is it possible to use both traditional and nontraditional methods simultaneously, or are they based on conflicting assumptions? Research in other contexts suggests that it is not always possible to use approaches that combine intrinsic and extrinsic measures.[79] Research conducted by various scholars[80] suggests that the interaction between intrinsic motivation and implicit behaviors is more complex than what the legal literature assumes.[81] My research is based on the assumption that govern-

[77] Waubert de Puiseau, B., Glöckner, A., & Towfigh, E.V. (2016). *Comparing and Integrating Theories of Law Obedience: Deterrence, Legitimacy, and Self-control.* Unpublished manuscript, University of Hagen, Hagen, Germany.

[78] See Fine et al., *supra* note 60, at 314.

[79] Feldman, Y. (2011). The complexity of disentangling intrinsic and extrinsic compliance motivations: Theoretical and empirical insights from the behavioral analysis of law. Wash. UJL & Pol'y, 35, 11.

[80] Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology*, 82(5), 835.

[81] Legault, L., Gutsell, J. N., & Inzlicht, M. (2011). Ironic effects of antiprejudice messages: How motivational interventions can reduce (but also increase) prejudice. *Psychological Science*, 22(12), 1472–1477, for evidence of backfire effects when attempting to change people's prejudice either explicitly or implicitly.

ments must invest in improving legitimacy and morality of law even in areas that seem to involve automatic behavior. As an alternative to focusing on individual behavior, the chapter examines the contexts in which governments, organizations, and individuals engage in improving their automatic ethical behavior. I also discuss the many trade-offs that the BE literature requires us to take into account. For example, as suggested earlier, is accountability good or bad? Does it undermine some biases, but allow others? How important is legitimacy in a dual-reasoning context? To what extent is it possible to treat both good and bad people using different enforcement mechanisms? Should we let people know about the existence of the nudge?[82]

Chapter 9: The Corruption of "Good People"

Individuals often feel that they are not being treated fairly by employers, public officials, or people whom they hire to attend to their best interests in various capacities such as lawyers, physicians, architects, and accountants. Professionals whom we trust to behave responsibly and to focus primarily on our interests (when we hire them), on the interest of the public (in the case of public officials), or on the workplace (in the case of the employer) turn out instead to be influenced by personal or competing institutional interests. For example, at the workplace, an individual who may be up for promotion may develop the impression that other candidates are more likely to know the decision makers personally and thus have an unfair advantage. In the areas of corruption and discrimination, the notion of trust in the system is highly important because it affects individuals' ability to trust the professionals who are expected to attend to their interest (lawyers, physicians, etc.), public officials (municipal officers), and hiring managers (in employment situations).

Often the potential deviation from objectivity and impartiality is relatively subtle, and the professional can easily deny or ignore it based on legitimate rationales. In hiring or promotion contexts and regarding the exercise of professional duties, subtle deviations usually occur in situations in which there is more than one legitimate choice, and therefore there is room for various interpretations of what is the right thing to do. In the presence of vagueness, people have greater room for self-deception and motivated reasoning, and we expect that good people are more likely to find ways to justify their bad behavior. Some argue that such deviations are beyond the reach of classical enforcement mechanisms. Yet the focus on subtle deviations from impartiality creates a rich ground for research on the interplay between legal theory, ethical decision making, and empirical analysis of the law.

Many deviations from impartiality and professional integrity are carried out without full awareness. But because the proportion of unethical behaviors that are

---

[82] This is investigated in Ariel Tikochiski's PhD research.

carried out with full awareness in any given context is unknown in advance, it is necessary to explore a hybrid approach to modifying human behavior that allows us to address both implicit and explicit violations, with limited cross-interference. Although soft interventions are criticized because of their threat to autonomy,[83] ethical nudges – which are aimed at curbing corruption and discrimination – generate a different type of policy trade-off, as explained earlier.

In this chapter, I pay special attention to answering the following theoretical and applied policy questions: What effect does group or institutional affiliation have on bias and impartiality? How do monetary incentives affect impartiality? To what extent are people aware that such effects can change their judgment as well as that of others, and how does it affect their trust in the integrity of these processes? To what extent do people believe in the efficacy of various traditional and BIT-related legal instruments? What legal interventions are most likely to be regarded as legitimate and improve public trust both in society and in the ability of the state to change human behavior in a sustainable way?

Focusing on implicit corruption, I compare the different intervention methods, examining their effectiveness in curbing people's behavior in situations of conflict of interest. I pay special attention to subtle conflicts of interest, where many good people may not recognize that there is something unethical about their behavior.[84] This focus is needed because, globally, countries experience some degree of corruption at most levels of government,[85] and corruption arises in a different form in the private sector (e.g., in the area of corporate governance).

Research on corruption and conflict of interest contains numerous examples of situations in which people who exhibit professional and moral responsibility have allowed their self-interest, admittedly without full awareness, to prevail over fulfilling their duties.[86] One of the most studied areas in this context is the conflict of interest of physicians who conduct clinical studies financed by pharmaceutical companies or who prescribe drugs based on their relationship to such

---

[83] See, for example, Arad, A., & Rubinstein, A. (2015). The people's perspective on libertarian-paternalistic policies and nudging and choice architecture: Ethical considerations; http://arielrubinstein.tau.ac.il/papers/LP.

[84] Here are a few examples of subtle conflict of interest situations: voting on the academic promotion of a friend, consulting for a firm that may compete in the future with one's current employer, when a civil servant treats an affluent entrepreneur with greater consideration than usual, when a physician performs a procedure because he is more comfortable with it but it may not be what the patient would most benefit from, or a lawyer rejecting a plea bargain or a settlement that is in the best interests of her client.

[85] For example, a 2012 report found that all European countries suffer from political corruption at some level. Mulcahy, S. (2012). *Money, politics, power: Corruption risks in Europe.* Berlin: Transparency International.

[86] There is a wealth of research on the prevalence of conflict of interest in almost every field. See, for example, Rodwin, *supra* note 46. See also Thompson, *supra* note 46.

companies.[87] Most clinicians do not think they are doing anything wrong when they prescribe a certain course of treatment to their patients while ignoring the subtle effects of competing interests. In many similar situations, most good people may believe that the option that promotes their self-interest is also the correct one. We can include in this group lawyers dealing with their clients, executives acting on behalf of shareholders, prosecutors making plea bargains, and academics deciding whether their colleagues should be promoted.

In the context of implicit corruption, psychological processes such as self-deception, elastic justification, moral disengagement, and motivated reasoning enable people to behave unethically without recognizing their wrongdoing.[88] As in research on prejudice and discrimination, a vast literature suggests that self-interest may influence people without their recognizing its effect on their behavior.[89] Moore et al.[90] showed that people truly believed their own biased judgments and had limited ability to recognize that their behavior was affected by self-interest.[91] These conclusions are also supported by the work of Gino et al.[92] and of Shalvi et al.[93] regarding honesty and by the work of Halali et al. regarding fairness.[94] Although the debate in the literature continues,[95] from an applied perspective, making the behavior of good people more ethical requires an understanding of implicit corruption and the fact that it is difficult to manage.[96] Various

---

[87] Rodwin, M. A. (2012). Conflicts of interest, institutional corruption, and pharma: An agenda for reform. *Journal of Law, Medicine & Ethics*, 3(40), 511–522.

[88] "Much of the problem with conflicts of interest is not intentional corruption but *unintentional bias*. Bias is widespread and is a problem even for well-meaning professionals": Cain, D. M., & Detsky, A. S. (2008). Everyone's a little bit biased (even physicians). *JAMA*, 299(24), 2893–2895; Loewenstein, G., Sah, S., & Cain, D. M. (2012). The unintended consequences of conflict of interest disclosure. *JAMA*, 307(7), 669–670.

[89] See Moore & Loewenstein, *supra* note 16.

[90] Moore, D. A., Tanlu, L., & Bazerman, M. H. (2010). Conflict of interest and the intrusion of bias. *Judgment and Decision Making*, 5(1), 37.

[91] "Private" evaluations were measured by giving participants incentives to be accurate in their predictions. See also Bazerman, M. H., & Sezer, O. (2016). Bounded awareness: Implications for ethical decision making. *Organizational Behavior and Human Decision Processes*, 136, 95–105, for a review of the mechanism responsible for limited awareness of unethicality.

[92] Gino, F., Schweitzer, M. E., Mead, N. L., & Ariely, D. (2011). Unable to resist temptation: How self-control depletion promotes unethical behavior. *Organizational Behavior and Human Decision Processes*, 115(2), 191–203.

[93] See Shalvi et al., *supra* note 13.

[94] Halali, E., Bereby-Meyer, Y., & Ockenfels, A. (2013). Is it all about the self? The effect of self-control depletion on ultimatum game proposers. *Frontiers in Human Neuroscience*, 7.

[95] For example, Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological Science*, 17(12), 1082–1089.

[96] See Lessig, L. (2011). *Republic, lost: How money corrupts congress – and a plan to stop it*. New York: Twelve. See also Feldman, Y., Gauthier, R., & Schuler, T. (2013). Curbing misconduct in the pharmaceutical industry: Insights from behavioral ethics and the behavioral approach to law. *Journal of Law, Medicine & Ethics*, 41(3), 620–628.

studies have shown that disclosure, which has been regarded as the ultimate solution for curbing corruption, does not work for implicit processes and can even have the reverse effect from that desired.[97]

Chapter 10: Discrimination by "Good" Employers

Employment discrimination is one of the most serious problems in labor markets worldwide and so has attracted more attention than other forms of discrimination (e.g., financial, residential). Anti-discrimination employment laws prohibit specific forms of employment discrimination – for example, that based on race, sex, religion, and age.[98] But usually these laws do not address each form of discrimination individually, nor do they take into account the different sociological and psychological mechanisms behind each form. In most countries, the legal approach is a general one, and similar remedies and prohibitions are applied to various forms of discrimination.

Research on the non-deliberative aspects of discrimination uses the insights of social and cognitive research in intergroup psychology,[99] which has focused on stereotyping processes.[100] Fiske's work is especially promising for legal scholars because it offers a nuanced and multidimensional approach to discrimination.[101] Psychological insights have been incorporated into the study of employment discrimination to a greater extent than in any other legal area.[102] Legal scholars, most

---

[97] See Cain et al., *supra* note 70.

[98] Civil Rights Act, Title VII (1964) prohibits employment discrimination on the basis of religion, race, and sex. The Age Discrimination in Employment Act of 1967 (ADEA) protects applicants and employees who are 40 years old and older from discrimination. Both laws are enforced by the US Equal Employment Opportunity Commission (EEOC). In the United States, the difference between the categories of discrimination is mostly due to constitutional reasons rather than behavioral ones.

[99] Tajfel, H. (Ed.). (2010). *Social identity and intergroup relations*. Cambridge: Cambridge University Press. See also Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. *Social Psychology of Intergroup Relations*, 33(47), 74.

[100] Much of the research in this field can be attributed to the writing of Fiske and collaborators, for example, Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878. See also Fiske, S. T. (2012). Managing ambivalent prejudices: Smart-but-cold and warm-but-dumb stereotypes. *The Annals of the American Academy of Political and Social Science*, 639(1), 33–48; Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology*, 72(2), 275; Fiske, S. T., Gilbert, D. T., & Lindzey, G. (Eds.). (2010). *Handbook of social psychology* (Vol. 2). Hoboken, NJ: John Wiley & Sons.

[101] Fiske, S. T. (2000). Stereotyping, prejudice, and discrimination at the seam between the centuries: Evolution, culture, mind, and brain. *European Journal of Social Psychology*, 30(3), 299–322.

[102] Bodensteiner, I. E. (2008). The implications of psychological research related to unconscious discrimination and implicit bias in proving intentional discrimination. *Missouri Law Review*, 73, 83, on circumstantial evidence; for the shift in the burden of proof in employment discrimination cases and for a broader account, see Pedersen, N. B. (2010). A legal framework for uncovering implicit bias. *University of Cincinnati Law Review*, 79, 97.

notably Krieger, have suggested that many biased employment decisions result not from discriminatory motivations but from a variety of unintentional categorization errors.[103] Considering the richness of behavioral findings on employment discrimination, the lack of responsiveness of the law to this knowledge is frustrating. Krieger and Fiske[104] discussed the outdated nature of US laws in this area from both jurisprudential and practical aspects. For example, the law requires showing intention and finding evidence for what has occurred at the stage of discrimination,[105] but it falls short of providing a comprehensive legal alternative to implicit discrimination.

Some legal scholars have acknowledged that most acts of discrimination are the product of a variety of unintentional errors.[106] As in most other types of research on non-deliberative choice, the recent literature on discrimination reveals the problems associated with automatic reasoning, but it offers almost no suggestions for the format of a new legal policy that would address both deliberative and non-

---

[103] Krieger, L. H. (1995). The content of our categories: A cognitive bias approach to discrimination and equal employment opportunity. *Stanford Law Review*, 47, 1161–1248; Krieger, L. H., & Fiske, S. T. (2006). Behavioral realism in employment discrimination law: Implicit bias and disparate treatment. *California Law Review*, 94(4), 997–1062. Krieger, L. H. (2000). Afterword: Socio-legal backlash. *Berkeley Journal of Employment and Labor Law*, 21(1), 476–520; Hart, M. (2004). Subjective decisionmaking and unconscious discrimination. *Alabama Law Review*, 56, 741; Cuddy, A. J., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology*, 40, 61–149; Uhlmann, E. L., Brescoll, V. L., & Machery, E. (2010). The motives underlying stereotype-based discrimination against members of stigmatized groups. *Social Justice Research*, 23(1), 1–16. See Agerström, J., & Rooth, D. O. (2011). The role of automatic obesity stereotypes in real hiring discrimination. *Journal of Applied Psychology*, 96(4), 790, in the context of overweight people.

[104] See Krieger & Fiske, *supra* note 103.

[105] A vast body of federal employment statutes provides protection against discrimination based on group membership. See Title VII of the Civil Rights Act of 1964, *supra* note 98, for example, prohibits discrimination based on race, color, religion, gender, or national origin. The Equal Pay Act prohibits employers from paying different wages based on the gender of the employee (but does not prohibit other discriminatory employment practices). The Age Discrimination in Employment Act prohibits employment discrimination on the basis of age. The Americans with Disabilities Act (1990) prohibits discrimination by employers based on physical or mental handicaps. In recent years, the legal situation is similar in Israeli courts, although the courts are aware of the existence of unconscious discrimination and try to objectively test for discrimination (H.C.J. 104/87 Dr. Neomi Nevo v. The National Labor Court, 1987, 09-9690 Ababa v. A.A 101 Group, 2013). Nevertheless, the vast majority of misconduct in this area never reaches the courts; this is true of employment discrimination in general, see Nelson, R. L., & Bridges, W. P. (1999). *Legalizing gender inequality: Courts, markets and unequal pay for women in America* (Vol. 16). New York: Cambridge University Press; and in Israel in particular: Sharon Rabin Margaliot (2000) The slippery case of age discrimination – how does one prove its existence? *Advoc* 44, 529, 537. See also Cerullo, C. (2013). Everyone's a little bit racist: Reconciling implicit bias and Title VII. *Fordham Law Review*, 82, 127, for a review of the treatment of implicit biases by US courts.

[106] See Krieger, *supra* note 103.

deliberative discriminatory behavior. For the most part, the law still looks for a smoking gun when identifying employers as committing acts of discrimination and prejudice.

Chapter 11: Summary and Conclusion

It is important both theoretically and practically to understand that there is no one-size-fits-all solution to incorporating considerations of the intrinsic versus extrinsic dynamic into government policies. It is difficult to predict the accumulated effect of these mechanisms without taking into account the context, and, in any case, the accuracy of predictions will always be limited. I therefore recommend that theoretical efforts be focused on creating a multidimensional taxonomy of contexts that elucidates the dynamics of intrinsic-extrinsic motivations.

First, what is the nature of the behavior we want to encourage?[107] It is important to take into account the behavior that the policy maker wishes to promote. One cannot excel in recycling or even in organ donation;[108] for the most part, policy makers care only about one's activity level and willingness to pay for engaging in recycling. In various other legal contexts, however, the quality of the behavior is more important. For example, in whistle-blowing or blood donation, it is less helpful to think about employees who do it for purely extrinsic reasons. In a legal context, where extra-role activity is desired, the cost of reducing intrinsic motivation increases, and one should be more cautious in introducing extrinsic motives.

Second, what proportion of the target population do we need to cooperate?[109] When the level of intrinsic motivation is heterogeneous, what proportion of the target population do we need to comply?

In the context of trade secrets, we need the cooperation of 100 percent of the target population, from those with the highest level of intrinsic motivation to those with the lowest. Therefore, the effect of reducing the intrinsic motivation of committed employees may be secondary to making sure that even those without intrinsic motivation remain loyal to their employers. In the case of whistle-blowing, the exact opposite is true, and we need the cooperation of only some of the employees to come forward when some illegal activity is taking place within the organization. Therefore, the policy maker can focus primarily on those who are high on intrinsic motivation.[110] For obvious reasons, we may not even want to provide incentives to those without intrinsic motivation because of fear of false reports by bounty hunters. Finally, in the context of recycling, we are interested in averaging, so that as many

---

[107] To date, I have presented data about three main types of activities (trade secrets, recycling, whistle-blowing): I use these three examples to help us think about the importance of being aware of legal contexts when policy makers attempt to decide how to provide incentives for certain behaviors without harming individuals' intrinsic motivations.

[108] But this is not the case with regard to blood donation.

[109] See Benabou & Tirole, *supra* note 75.

[110] This argument is obviously oversimplified, and fine-tuning is highly needed here.

people as possible recycle as much as possible. In this situation, we have no preference for either high or low intrinsically motivated individuals; therefore, the balancing consideration for the policy maker is whether or not to use extrinsic motivation and, if yes, determining which types of incentives to use.

Third, how important is it to think that others are being motivated by intrinsic motives? People are biased in their perceptions of what others are doing and for which reasons. It is clear, however, that the effect of knowing why other people do what they do is different depending on the context, the nature of the relationship, the level of reciprocity, the importance of others' motivation to one's evaluation of its authenticity, and more. Presumably, the closer the behavior is to areas where one would expect identity-related factors to be dominant, the greater the damage is to the other from viewing one's motivation as being extrinsically motivated. In commercial contexts, we are less likely to find that extrinsic motivation harms perceptions of authenticity in the behavior of others. A relatively straightforward aspect we may want to consider is the visibility of the behavior and the ability to measure both its quantity and quality (recycling in houses vs. loyalty to an employer in keeping proprietary information secret). It is safe to assume that with more visible and measureable behavior, the policy maker should care less about harming intrinsic motivation whose main advantage is its limited dependence on external measurement. Thinking about these context dimensions could lead the policy maker to focus efforts on protecting intrinsic motivation in the most suitable contexts.

In the next section of the conclusion chapter, I examine the ability of law to change people's implicit tendency to behave unethically, relating this discussion to the research conducted by scholars such as Devine and Inzlicht on the ability to change people's intrinsic tendency to rely on stereotypes. Furthermore, because the focus of this book is on legal enforcement, I examine whether existing mechanisms through which the law could change the social meaning of certain behaviors could end up also affecting the unethical behavior of people that is based on implicit processes.

The concluding section of this chapter summarizes the many important questions we have discussed in the book and suggests that they remain open to further research. Because some of the most important interactions between psychology and law have not yet been studied, the book cannot be expected to resolve all these issues or even address them. In this section, I focus on emerging research directions. First, I present the set of jurisprudential questions that should be addressed more extensively, such as free will, autonomy, variations among people, equality, and the role of law relative to morality. As long as these questions are being addressed mainly from a legal enforcement perspective, without extensive involvement by scholars of law and philosophy, the ability to change the field remains limited. Second, I present a series of research questions on enforcement mechanisms. For example, given what we know about good people, should we change the

desired standard of behavior? How much information should the state collect on people's attitudes and preferences? Should states impose standards of behavior whereby organizations are actively required to actively engage in regulating situations? Should people's and organizations' failure to adopt certain situational procedures to prevent unethicality put them in some legal responsibility by omission?