## 

Contents



[Intermediate summary 2](#_Toc486936192)

[Differentiated vs. integrated, or sequential regulatory approach 4](#_Toc486936193)

[Shifting the focus of government intervention from ex post to ex ante 5](#_Toc486936194)

[Can we create good people: The Fall and Rise of Intrinsic Motivation and Morality 5](#_Toc486936195)

[Between motivation and cognition 6](#_Toc486936196)

[Changing people’s non-deliberative choices, through external regulatory interventions. 7](#_Toc486936197)

[Intrinsic Motivation and Prevention of Ethical Biases 9](#_Toc486936198)

[Motivational and cognitive training and the reduction in ethical biases 11](#_Toc486936199)

[Taxonomy as a way to modify the one fits all policy 13](#_Toc486936200)

[The nature of the behavior 14](#_Toc486936201)

[Cooperation of what proportion of the target population do we need for the policy to work? 14](#_Toc486936202)

[Is there a Moral Consensus with regard to the behavior in question? 15](#_Toc486936203)

[The Needed Durability of the Behavioral Change 17](#_Toc486936204)

[Costs of Mistakes 18](#_Toc486936205)

[Conclusion 18](#_Toc486936206)

[Revising enforcement strategies 21](#_Toc486936207)

[Smarter usage of uncertainty 21](#_Toc486936208)

[A statistical approach 22](#_Toc486936209)

[The need to develop and integrative approach through data gathering 23](#_Toc486936210)

[According to most views in ethical context, self-interest is more intuitive than fairness concerns. For various motivational and cognitive processes, People don’t (always/never?) know that they are about to break the law. Procedural justice and dual reasoning 23](#_Toc486936211)

[Kith van den boss seems to suggest that 23](#_Toc486936212)

[Limitations 24](#_Toc486936213)

[The future of behavioral ethics in law 26](#_Toc486936214)

[The book leaves us with many unanswered questions which future research should discuss and refine: How blind is a blind spot from a legal perspective of responsibility? Can we combine traditional methods with the new BIT-type methods? What knowledge is needed to offer legal policy makers a formula for optimally balancing traditional intervention methods with non-traditional ones? Can we combine traditional and nudge-like interventions without harming either approach? What values are more important when attempting to increase the efficacy and legitimacy of nudges (e.g., making people aware of the nudges)? Can we know ex-ante in what mode of reasoning people will be when making a decision about the law? Can we know in advance what is the effect that similar interventions on good and bad people would have in every policy context? 26](#_Toc486936215)

[[*This part is very preliminary*] 27](#_Toc486936216)

[Removed stuff 27](#_Toc486936217)

[Further research 27](#_Toc486936218)

## Intermediate summary

In this book, I created the infrastructure for the neglected area of research on human behavior in legal theory and practice – behavioral ethics. People’s inability to fully capture the social, legal, and moral meaning of their behavior and the variation in people’s cognition and motivation toward the law challenge the main ideas behind both legal liability and legal enforcement. In a world that assumes there are only calculative individuals, the set of tools and the predictions about human behavior are relatively clear-cut. Yet the good people typology introduces complications and complexity. It does not replace the traditional model, but increases the number of models that legal policy makers need to deal with simultaneously.

This book’s basic premise is that there are three types of people: the traditional “calculated” wrongdoers; the genuinely moral individuals whose wrongdoing is based on a blind spot; and the most challenging group of “situational” wrongdoers, those who use various social and situational cues to justify unethicality. The bad/calculated people are not about to disappear, but are just one of at least three types of people whose wrongdoing should be regulated both ex-ante as well as ex-post.[[1]](#footnote-1) We also know that there are individual differences and variation within each group, as discussed in Chapter \_. Complicating the situation further is that many in the largest group[[2]](#footnote-2) of "good people" might still be somewhat aware of their wrongdoing. As shown in Chapters 2 and 5 many of the cognitive mechanisms that people use make them somewhat of the processes through which they justify unethicality. The variation both in the level of awareness of wrongdoing and mechanisms used to allow people to break the law while still feeling that they are normative individuals makes it even more difficult to regulate “good people.”[[3]](#footnote-3)

In this final chapter, I present concluding thoughts on the interaction between behavioral ethics and law. I examine the important role of intrinsic motivation in a “good people world,” given the overwhelming evidence of its ability to change both their perception and behaviors.[[4]](#footnote-4) I then present a taxonomy that would help the state organize its regulatory and enforcement efforts. I conclude with a discussion of limitations in BE research, unanswered questions, and research efforts needed to answer them, so that states will be able to more effectively regulate the behavior of both good and bad people.

Research into the dual reasoning systems, the crowding-out effect, and intrinsic versus extrinsic motivation makes clear that people react to the law in more ways than previously assumed. This book therefore advocates an integrated approach that incorporates both nontraditional and traditional methods of regulation and enforcement.. This approach reflects a change not only in the mix of interventions used but also in their design and sensitivity

Yet, even an integrated approach finds it difficult to determine ex-ante who are the people who are likely not to comply with the law. We have suggested various methods to account for this variation in the population: acoustic separation, use of a taxonomy, and accounting for the level of intrinsic motivation toward a specific legal doctrine. For example, legal instruments, such as regulations and contracts, can be designed in ways that differentiate between the motivations of good and bad people to meet their obligations;[[5]](#footnote-5) differences in cultures can be addressed when attempting to change tax compliance norms. Designing rules in a legitimate way and creating effective deterrence methods can also help reinforce traditional mechanisms of enforcement, which will affect both good and bad people. Ex post, the research paradigm can help courts and other enforcement agencies understand the severity and intentionality of various misconducts.

Following Ayres and Braithwaite responsive regulation pyramid of regulation approach[[6]](#footnote-6)

According to Tyler, it is possible to design interventions based on people’s intrinsic motivation, beginning with those that target considerations of morality, fairness, and social values and then moving to a harsh approach to the minority of people who are more calculated wrongdoers. This sequential move from soft to hard regulation may align with a move from nontraditional and situational enforcement, which focuses on good people, to traditional enforcement, which is always in the background but will come into play when nontraditional enforcement fails. However, while the approach of Tyler and that of Ayres and Braithwaite holds that soft regulation is mostly preferable to hard regulation, we recognize that nontraditional means such as nudges suffer from many limitations that make them inferior to traditional enforcement methods that involve deliberation.

The rest of the chapter focuses on two main strategies that legal policy makers can adopt to address variation among people: reducing the impact of ethical biases by changing people’s motivation and using taxonomies that offer a set of considerations to take into account when fitting the right type of intervention to a specific behavior.

## 

## <A>Can We Create Good People: The Fall and Rise of Intrinsic Motivation and Morality

Earlier chapters examined the role of morality in legal compliance and how to increase levels of intrinsic motivation to obey the law. This section draws attention to another aspect of intrinsic motivation: its ability to reduce nondeliberative biases and thus make the automatic system more ethical. The goal of legal policy can thus be reformulated to create a situation in which people are intrinsically motivated not to be cognitively biased to their own unethicality.

There are a growing number of studies that demonstrate that by changing people’s motivation, we are able to change their nondeliberate wrongdoing.[[7]](#footnote-7) The best-known research is by Divine,[[8]](#footnote-8) who has studied automatic and controlled components of stereotypes and prejudice and the ability to change those automatic components. She and Inzlicht, another noted researcher, found that people can be motivated to overcome their unethical biases, but it is unclear to what extent this change is durable over time. I am not yet aware of a study that has compared, for example, the efficacy of nudges vs. intrinsic motivation in changing people’s ethical behavior in the long run. However, it seems from the literature that motivation can sometimes change implicit biases without any cognitive tricks, mostly through changing people’s intentions.

This accords with another of our basic premises —that there has been too much attention to the role of cognitive mechanisms in influencing behavior and too little focus on the important role of intrinsic motivation. What we learn from scholars such as Devine and Inzlicht is that intrinsic motivation matters for behaviors with limited cognition. Even under legal policy that is sensitive to nondeliberative choice and even with various nudges that communicate directly with people’s automatic systems, motivation plays an important part in regulating good people and facilitating their compliance, rather than just punishing them for their unethicality.

## 

## <B>Motivational and Cognitive Training in Reducing Ethical Biases

Bazerman and Tenbrunsel (2011) have suggested various techniques for people to become aware of their limited ability to recognize their ethical failings and so reduce their prejudices and in general behave more ethically. For example, focusing on the “should” rather than the “want” self, Bazerman and Gino suggested that people should consider beforehand what “want” desires might come into play at moments of decision making. They can then be better prepared to resist these desires and instead implement a decision based on their ethically sound “should” preferences. Bazerman and Gino use the example of a common interview question asked of job applicants: What pay has a competing employer offered you? The “want” self is likely to wish to inflate the number, encouraging the applicant to lie to the potential employer. By anticipating this, one can come up with a more acceptable answer, such as “I’m not comfortable sharing that information,” which serves the applicant’s self-interest but also does not violate moral rules. This “should” vs. “want” approach recognizes areas in which people might cut corners and tries to anticipate those situations in advance.

Similarly, in a series of experiments, Devine et al. (2012) showed that it is possible to train people to be less affected by their implicit racial biases and this change is long-lasting. At a minimum, these experiments support the claim that even if we are unaware of the automatic components of our behaviors, we may be able to control them after participating in training. In the field of social cognition, which is mostly related to employment discrimination, there seems to be consensus that stereotypes can be altered with various training techniques.[[9]](#footnote-11) Bartlett suggests that for people to be able to change their implicit associations, they need to feel that they have autonomy and free choice or else any external intervention might backfire.[[10]](#footnote-12) The argument goes that, using some form of legal intervention or deterrence, which by definition calls for some restriction on people’s autonomy, might negatively affect the ability of people to reduce implicit bias.

## <B>Between Motivation and Cognition

As mentioned earlier, the role of intrinsic motivation is related to the broader discussion that lies in the heart of behavioral ethics: the interaction between motivation and cognition, particularly how people’s motivation toward the law affects the likelihood of transgressions happening with limited awareness. The literature reviewed in Chapter 2 suggests the important role of enhanced self-control, which is mostly associated with system 2 reasoning, in curbing unethicality. Thus, heightening the awareness of people that they need to behave ethically might force system 2 to kick in. However, to the extent that we view this awareness as the product not only of the situation but also of the motivation to avoid doing harm, policy makers should do more than target people’s cognitive awareness: they need to also make people want to avoid engage in that wrongdoing. Thus, whether people are “good” or “bad” seems to be related both to motivation and cognition. Furthermore the distinction between the two types of good people makes the interaction between motivation and cognition even more complex. The genuinely moral individual will only do bad things under conditions of limited cognition. However, situational wrongdoers are somewhat aware that what they are doing is bad, but their motivation is too weak to prohibit them from doing it. Seeing that everyone is parking illegally mostly reduces their motivation to follow the rules and increases their ability to view themselves as good people while parking illegally despite knowing that it is illegal. Clearly, when such behavior becomes a habit there will be more and more reliance on automatic reasoning; hence with limited motivation to comply with the law, we might see an increase in people engaging in misconduct that does not cause any moral dilemmas for them. Such processes are hard to measure in classical short-term psychological experiments, but are widely recognized in the sociological research on internalization.[[11]](#footnote-13)



## <B>The Perils of Changing People’s Non-Deliberative Choices Through External Regulatory Interventions

## 

Changing people’s level of intrinsic motivation may seem to be futile, as external motivation techniques such as incentives may crowd out and inhibit altruistic motivations. However, some studies do support the idea that extrinsic motivation can increase intrinsic motivation. The relationship between regulation and intrinsic motivation is a complicated one and is affected by several variables. For example, my work with Smith and Boussalis explored the interrelations between specificity, fairness, and monitoring. As described in Chapter \_\_, specificity in itself can interact both with morality and deterrence in any given context. This line of research is supported by work on the expressive function of the law, described in Chapter \_\_ , with a focus on the ability of law in the areas of trade secrets, environmental law and file sharing to change people’s perception of what is moral.[[12]](#footnote-14)

As described in more detail in Chapter \_ on the pluralistic function of law, the crowding-out literature generally finds that extrinsic motivation undermines intrinsic motivation. Fehr (2002) argued that when people attribute their behavior to external rewards, they discount their moral incentives for their behavior, thereby lowering the apparent effect of intrinsic motivation. For example, paying people to donate blood causes donors to view the donation as a transaction, rather than a charitable act, eroding altruistic blood donations. Similarly, in a series of lab-based experiments, Deci and colleagues[[13]](#footnote-15),[[14]](#footnote-16) argued that “tangible rewards tend to have a substantially negative effect on intrinsic motivation.” In a similar vein, Marshall and Harrison[[15]](#footnote-17) suggested that the use of incentives can damage self-esteem, resulting in the perception that professionalism is no longer valued.

Thus, external interventions can sometimes have reverse effects.[[16]](#footnote-18) Inzlicht and colleagues examined the effect of motivation (internal-autonomous vs. external controlled) on people with explicit prejudice. They showed that intrinsic motivation techniques focusing on increasing autonomy were more successful than using control to motivate prejudice reduction. In a second study that replicated the same type of motivations, but by using priming, a similar pattern emerged, where controlled motivation had what they called an “ironic” effect on prejudice.

Yet a more nuanced view holds that the effect of incentives is not linear, but rather that “intermediate” payouts have a disproportionately high crowding-out effect compared with low or high payout levels. Tenbrunsel and Messick[[17]](#footnote-19) argued that a weak system of sanctions produces worse results than having no sanctions at all. By introducing a sanctioning system, the principal changes the evaluation of the problem from an ethical dilemma to a business decision, consequently shifting individual considerations away from ethics and toward self-interest.[[18]](#footnote-20) Research done by Devine[[19]](#footnote-21) has shown that explicit race bias was moderated by internal motivation to respond without prejudice, whereas implicit race bias was moderated by the interaction of internal and external motivation to respond without prejudice. Specifically she found that high internal, low external participants exhibited lower levels of implicit race bias relative to other participants. She developed concept scales related to the existence of individual difference in the internal level of motivation to avoid prejudice.[[20]](#footnote-22) Similarly, Bartlett found that, when people who do not have the internal motivation to avoid racial bias are being threatened or coerced, they are unable to change their evaluation in the same way as people who have a stronger commitment to avoid racial bias. Her work highlights the challenge of motivating people to change without them feeling coerced.



The variation in the efforts of people with high and low motivation to avoid prejudice suggests the legitimacy of dealing with people as having different moral responsibilities. This is particularly salient, given the research showing that stereotypes can be altered and changed with training. Our conclusion in earlier chapters that people do not choose to do wrong, but rather interpret the situation in a way that will allow themselves to feel moral, suggests that we need to revisit their moral judgment analysis. People do not make moral decisions in a vacuum. Hence, the focus in behavioral choice is not on making morality salient, but on making it clearer to people that they have a position that might affect their moral judgment.

Furthermore, from a policy-making perspective, the relevant question might not be whether intrinsic motivation can increase compliance or performance, but whether and when intrinsic motivation outperforms extrinsic motivation in securing desirable behavior. In contexts in which intrinsic motivation is more successful in securing compliance, policy makers must make sure that the standard extrinsic motivators provided by law (e.g., sanctions and incentives) do not undercut intrinsic motivations.



Is it also the case that we can make people become more ethical? When an individual tries to reduce racial bias, he or she has to address only one dimension of character and behavior. Reducing racial bias thus seems to be simpler than situations regarding ethical behavior, where people have to choose from a number of options and have less clarity of what is the right thing to do. Acting ethically is related to many dimensions: honesty, rule compliance, cooperation, and many more attributes that make changing intrinsic motivation so challenging.

Even though increasing intrinsic motivation is likely to positively affect behavior, extrinsic motivation can be effective in discouraging wrongdoing. In addition, we still need to address the behavior of those with low intrinsic motivation to act ethically. We fully support education and training as important tools in making people more likely to obey laws, but to present them as the only tools will neither be effective nor practical.

In conclusion of this part, it seems that despite the limitations of motivation’s effect on ethical behavior, states should recognize that education and training can be effective tools to enable people to deal with many ethical problems. Yet people will only make a lasting change in their behavior when they change what they think is the right thing to do. In a sense, nudges are like rules that ensure that people do not engage in self-deception, but in and of themselves can never be a permanent solution that will lead to changing people’s intrinsic motivation. Finding ways to change people’s intrinsic motivation in a noncoercive way is likely to help reduce people’s ethical biases.

# <A>Taxonomy as a Tool to Modify the One-Size-Fits-All Policy

There are two main approaches to designing interventions: an integrated approach, where all means are used simultaneously, and the differentiated approach, in which a different approach is adopted for each situation. The most relevant conceptual framework for the differentiated approach is a taxonomy that includes the most relevant factors when attempting to implement and enforce a policy initiative: the type of regulated behavior, target population, the cost of enforcement and of noncompliance, and the proportion of the population whose cooperation is required for successful implementation of the mechanism.

The taxonomy presented here highlights the areas in which greater scrutiny of the ethical behavior of individuals should be exercised. It takes into account both traditional factors, such as the level of monitoring, intrinsic loyalty toward the organization, and compliance with prevailing norms, as well as notraditional factors: the type of individual we are attempting to regulate, the moral “wiggle room” of the situation, the relative availability of excuses for justifying behavior, and whether the task can be accomplished without cognitive focus. As suggested in earlier chapters, various policy trade-offs will emerge because each intervention has its own pros and cons. Trade-offs and dilemmas include the following: should there be public discourse on nudges, should we worry more when people make decisions in a group context, should we revise deterrence to deal with nondeliberative effects, and how can we affect “good” people without harming the credibility of state treatment of the “bad” ones.

## <B>What Is the Nature of the Behavior?

The key component is the behavior the policy maker wishes to promote, and its nature will determine the relative importance of intrinsic motivation. Why one engages in recycling or organ donation is not important[[21]](#footnote-23); for the most part, policy makers only care only about the activity level and willingness to pay. In other contexts, however, such as whistleblowing or even blood donation, intrinsic motivation seems to play a larger role. Furthermore, in legal contexts, where “extra-role” activity is desired, the cost of harming intrinsic motivation increases and one should be more cautious in introducing extrinsic motives. Where there is a high cost of enforcement or a great need for sustainability of behavior, there may need to be a stronger reliance on deliberation and preventing illegality.



## <B>What Proportion of the Target Population Needs to Cooperate for the Policy to Work?[[22]](#footnote-24)

Another important dimension is what proportion of the target population should be targeted when the level of intrinsic motivation is heterogeneous.[[23]](#footnote-25)

In earlier research, I focused empirically on the ability to change the behavior of people in three types of activities: recycling and file sharing, which I put in a category of “the more the merrier”; whistleblowing, which I put in the category of “it only takes one to help”; and divulgence of trade secrets, which can be put in the category of “it only takes one to harm.” These three examples help us think about the importance of being aware of legal contexts when determining how to change the behavior of a given population with regard to a given doctrine.

In the area of trade secrets,[[24]](#footnote-26) everyone needs to be motivated in order for the secret not to be disclosed, because even if only a very few people are not affected by the legal instrument, it may be futile to keep company knowledge proprietary. Clearly the value of the trade secret diminishes when more than a handful of unauthorized people know about it. In this context, we have to focus on the lowest common denominator, meaning that those with the lowest level of intrinsic motivation to be loyal to the company should be the focus of the regulation. Since variation in motivations is likely to increase the chance of making mistakes and mistakes are costly in such direction, a greater emphasis should be given to making sure that at the least we get minimal compliance by all employees. The price of harming the intrinsic motivation of committed employees might be secondary to making sure that even those without intrinsic motivation will be loyal to their employers.

The whistleblowing context is the exact opposite, where we only need the cooperation of some of the employees who will go forward when some illegal activity occurs within the organization. Therefore we mainly care about those who are high on intrinsic motivation[[25]](#footnote-27) and may not even want to incentivize those without intrinsic motivation due to a fear of generating false reports from bounty hunters.

Finally, in the context of recycling,[[26]](#footnote-28) and also file sharing,[[27]](#footnote-29) we are interested in averaging or in a situation where as many people as possible will recycle as much as possible. Outcomes are important, but they are long term and aggregate. Although the ultimate goal is to move as many people as possible to environmentally responsible behaviors, private noncompliance is not very costly. , must balance In this context making a few mistakes in targeting people’s motivations is not costly since the effort is to increase the average level of recycling.

## <B>Is There a Moral Consensus with Regard to the Behavior in Question?

Another important dimension derived rom the research on intrinsic versus extrinsic motivation is the extent to which the behavior that the state is attempting to regulate is within the moral consensus of the population.[[28]](#footnote-31) When there is a moral consensus and a population is likely to have a high level of intrinsic motivation to comply, the policy maker could rely on noncoercive explicit measures or an increase in the proportion of softer types of implicit intervention.

In my work with Lobel,[[29]](#footnote-32) we found that the level of moral consensus had an interesting effect on regulation: the greater the perception of severity of the misconduct, the less important was the choice of regulatory mechanism. In our analysis, we used severity of the misconduct as a proxy for internal motivation.[[30]](#footnote-33) In the group of participants who viewed the illegality as highly offensive, and hence had high levels of internal motivation for reporting, the type of mechanism available to them to do so was largely irrelevant. Respondents expected their own reporting levels and those others to remain consistently high across all categories of legal mechanisms. Thus, i

However, when illegalities witnessed by potential enforcers were perceived as less severe, the use of high rewards and fines produced considerably higher levels of reporting than the use of low rewards. These findings suggest the importance of legal mechanism selection in instances where individuals do not have an ethical stake in compliance. In such cases, triggering external motivation through regulatory policy takes on a far greater role in promoting reporting activity. Therefore, regulatory agencies may consider providing high monetary rewards when the goal is to incentivize reports in contexts that evoke less moral outrage, such as tax evasion.

The same argument seems to apply to the good/bad people dichotomy: a greater focus on extrinsic measures should be used in areas where more percentages of the population are unlikely to explicitly be interested in pursuing the behavior. That is why crime prevention focuses on the “bad’ person, even though the vast majority of the population will be restrained from committing a crime by their moral constraints. However, the harm caused by even one person committing a crime is so high that we focus on harsh, external methods.

## <B> How Long Does the Behavior Change Need to Last?

Whether the behavior only occurs once or is repeated is another important consideration. For example, in encouraging enrollment in a pension plan, the focus is on the initial decision; getting people to participate (outcome being the dominant focus) is important. Once they have made their choice, people are less likely to reverse it, and sustainability is less important than in areas like health or nutrition, where choices need to be reaffirmed on a daily basis. Similarly, the focus on expressive versus invisible law is also dependent on context. In areas where the expertise of the state as well as the moral or consensual bases are high, highlighting the expressive law might outweigh the costs associated with informing people that that the choice architecture presented to them is based on law. In social issues in which preferences for a deliberative process are strong and existing legal solutions are contested, more weight should be given to process. Focusing on trust may be more important in areas that are difficult to monitor, whereas focusing on directed regulation is desirable when enforcement costs are relatively low.

## <B> What Is the Cost of Noncompliance?

In some situations, the costs of noncompliance are disproportionately large relative to the benefits of performing intrinsically. Since variation in motivations is likely to increase the chance of regulation being ineffective and mistakes are costly, in-depth analysis of the level of desirable compliance and its counter-costs becomes crucial. For example, in the context of trade secrets, one egregious leak may be detrimental to a company,[[31]](#footnote-34) while with many environmental protections, a failure of one individual to recycle is not consequential because what are important are long-term and aggregate behaviors. Some private noncompliance is thus not very costly.

# <A> Conclusion

Although we traditionally think that the law’s main purpose is to protect us from the bad people who breach contracts, make a nuisance of themselves in public places, engage in corrupt behaviors, steal intellectual property, use damaging speech on Facebook, cut in lines, or use nepotism to advance their careers, this book suggests that it is the good people whom we need to worry about. Behavioral ethics claims that the treatment of the good people who commit many of these bad behaviors is the neglected task of the law.

In contrast to the traditional approach that argues for a focus on optimism bias to understand why people become involved in the legal system, the behavioral ethics approach suggests that people go to court because they fail to fully capture how problematic is their behavior, due to the functioning of various mechanisms that prevent them from recognizing their wrongdoing. Existing enforcement strategies are not suitable for these types of people, and so new tools and modifications of traditional tools are essential.

In Chapter 2 we showed that a set of both deliberate and nondeliberate mechanisms such as moral forgetting and motivated blindness prevent people from recognizing the wrongdoing in their behavior and their own unethicality. Many legal disputes are not a product of people miscalculating their winning chances, but rather of their inability to recognize in an objective way their unethicality.

Chapters 3–5 addressed how to expand the regulatory toolbox, focusing on issues such as situational design, behavioral incentives, 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 will change people’s ex-ante calculations will not work because most people are likely not to be aware why they behaved in a certain way in the first place. In terms of fairness, we argued for the importance of designing policies that make it difficult for people to interpret fairness in a self-serving way. With regard to social norms, there is a need to provide people with accurate information on the nature of the norm and on the percentage of people who hold a certain norm. Incentives need to be sensitive to their crowding out 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 discussed the two types of good people. One type does not see reality as it is, because of Bazerman’s blind spot, Haidt’s approach to Morality,[[32]](#footnote-35) or Balcetis’s motivated seeing,[[33]](#footnote-36) which rely on nondeliberative mechanisms. In that camp of the morally blind people belong those who engage in implicit job discrimination or in implicit corruption in subtle conflict-of-interest situations. The second type of good person people knows that what they do is impermissible, but they find various rationales that they can use to allow themselves to do bad things and still feel great about themselves. The related work of Bandura on moral disengagement or Shalvie’s work on justified dishonesty shows that this is mostly a deliberate process. In the camp of the justifiers we can find those committing various parking violations, cutting in lines, or using contacts to get a job. In this chapter we also analyzed numerous relevant individual differences scales, such as moral identity and social value orientation.

Chapter 7 on the pluralistic effect of law and Chapter 8 on the tradeoffs between the different likely effects of laws recommended how policy makers could balance the effects of the different aspects of the law, on different people, with regard to different behaviors.

This new approach to law was applied in Chapters 9 and 10 to corruption and employment discrimination; these two case studies showed how to create an effective balance between different regulatory tools and the need to address different types of population with different states of mind toward the law.

<A>New Directions in Behavioral Ethics

<B>Situational Design

Given the fact that many individuals make choices without full awareness, it might make sense to place more of the responsibility for wrongdoing that occurs in organizational contexts on the organization itself. Doing so will incentivize the organization to design the situation in a way that will facilitate compliance and discourage wrongdoing. This shift in responsibility should be reinforced by state regulation and legal policy, which will focus on making it hard for people to find rationales for noncompliance.

### <B>Smarter Use of Uncertainty

Legal ambiguity has varying effects both on the ability to understand the meaning of various concepts in laws, contracts, and organisational codes and on subsequent decision making. Haisley and Weber,[[34]](#footnote-38) for example, found that people prefer ambiguous risks when such ambiguity allows them to justify their unfair behavior. Dana et al.[[35]](#footnote-39) found that people are less generous in situations in which they can use moral ambiguity to explain their selfish behavior. Similarly, Hsee[[36]](#footnote-40) found evidence that people make choices that satisfy their preferences at the cost of not completing an assigned goal, if they can exploit existing ambiguity about what decision can be considered as having achieved the assigned goal.

## <B>An Aggregate Approach

Given the difficulty of determining individuals’ awareness of the unethical nature of their decisions, BE suggests focusing on aggregating people’s decisions as proof of wrongdoing. This view can be understood by analogy from work in the area of employment discrimination,[[37]](#footnote-41)

where the inability to penetrate people’s minds (Krieger 1995) has led, in some cases, to an aggregated approach to evaluating decision making. Beginning with *Griggs v. Duke Power Co.,* (1971) the U.S. Supreme Court has recognized that, although it is not mandated that the workforce of a company should replicate the composition of the general population, statistical disparity between the two can be used as compelling evidence of employment discrimination under a disparate impact theory. According to this theory, even if it is impossible to prove that the employer intended to treat candidates differently, the fact that the employer used criteria that resulted in discrimination against a class of individuals is sufficient to establish illegitimate discrimination (Shoben 1983).



Similarly, in the realm of bounded ethicality, rather than assuming bad intentions where we cannot prove them, it may be possible to collect data on ethical decisions over time and create criteria to be applied if the aggregation of behaviors indicates that one should have been aware of the negative effect of one’s actions. For example, misuse of office supplies or the improper acceptance of gifts may be considered misbehavior even if any one instance of such conduct is merely questionable. A sufficient number of marginal instances can warrant sanctioning, regardless of the actor’s intent. Important jurisprudential work needs to be done to justify increasing one’s responsibility for one event based merely on the fact that it has been repeated. However, given the difficulty of determining responsibility for isolated events and the ability of System 2 to predict the likelihood that such unethicality will recur, a solution of this type may be necessary and appropriate.



## <B>Data Collection to Support the Integrative Approach

An Integrative approach is a complex approach. It involves a mix of both explicit and implicit interventions, as well as external and internal ones. It also recognizes that each law may affect behavior in more than one way. For example, the fact that a specific law focuses on deterrence does not mean that it does not have other impacts. Almost all legal instruments have the potential for an inadvertent effect, either because of a crowding-out effect, backlash, or targeting one population on the expense at others. Our discussion of behavioral trade-offs suggested that each intervention has vulnerabilities, and a broadly defined cost-benefit analysis needs to take place to assess each one’s effectiveness. A comprehensive and continuing data collection effort, as well as pilot initiatives, is essential to ensuring the most appropriate fit between law and each context.

## <A>Limitations of Beahvioral Ethics

Behavioral ethics makes an important contribution to the BLE literature, supporting the view that the self-serving effects of motivation on cognition allow people to do harm when it serves their self-interest without feeling guilty about their actions. The BE literature highlights biases that prevent people from understanding that their behaviors are self-interested and unethical. Uncovering these biases is especially important because society is harmed by nondeliberate bad deeds, which are not yet targeted by state interventions.

Yet, the ability of the current BE literature to recommend how to translate its findings into law is limited. Many important aspects are still being debated in the literature, both theoretically and methodologically. Within the concept of bounded ethicality, the interrelations between automaticity, awareness, and controllability are still the subject of controversy, and potential solutions are elusive. Furthermore, we know more about the effect of System 1 on System 2—which is of descriptive interest—than about the effect of System 2 on System 1—which is of greater normative interest. Even one of the most momentous questions for the law—are we intuitively good or bad—seems to be affected by context more than was previously assumed. In addition, the degree in which the good/bad aligns with individual differences is more limited than anticipated and seems to vary across different behaviors and contexts.

The lack of consensus among researchers on basic concepts limit the ability to base policies on BE findings. For example, would giving people more or less time to make behavioral choices increase or decrease the likelihood that they would engage in wrongdoing?[[38]](#footnote-42) How can we answer that question, if we do not yet know the meaning of methods such as time pressure (how much time) and ego depletion (how hard)? Finally, BE’s call for a differentiated treatment of good and bad people rests on the assumption that good people are genuinely unable to recognize the wrongness of their behavior. However, there is very little research on that topic, which limits the ability to remove responsibility from those good people, because it is very hard to rule out the option of them playing it dumb to avoid punishment.

Methodological issues also limit the ability of policy makers to translate the knowledge accumulated in behavioral ethics into public policy. from lab experiments,--situations studieds at workwhere The unethical behavior studied in the lab often involves people reporting a number from dice that they throw or the number of questions they answer or what number they see on a computer screen. In the real world, unethical behaviors are much more complex and are done without any deliberation, which raises doubts on the external validity of some of the lab studies involving unethical behavior. In addition, data are collected from laboratory experiments that show only the short-term effects of various ethical manipulations, such as , and for the most part, the law is more interested in the long-term effects of these practices. Most of the research in BE has been in a few areas of micromanagement, which is not entirely suitable for incorporation into legal theory and policymaking. Finally, debate continues on the validity of methods such as IAT and fMRI regarding the consistency of measures and predictability of behavior across situations.

## The book leaves us with many unanswered questions that future research should discuss and refine: How blind is a blind spot from a legal perspective of responsibility? Can we combine traditional methods with the new BIT-type methods? What knowledge is needed to offer legal policy makers a formula for optimally balancing traditional intervention methods with nontraditional ones? Can we combine traditional and nudge-like interventions without harming either approach? What values are more important when attempting to increase the efficacy and legitimacy of nudges (e.g., making people aware of the nudges)? Can we know ex-ante in what mode of reasoning people will be when making a decision about the law? Can we know in advance what is the effect that similar interventions would have on good and bad people in every policy context?

More specifically, who is the real person: the 2-second individual or the 20-second individual? The one who acts without deliberation or the one who is calculative? Given research that shows people can change their stereotypes and become less biased through training, should we impose a duty on people to engage in such ethical courses and impose liabilities on those who do not pass them? Should we impose costs on organizations that do not design their environments to facilitate ethical behavior?

The main theoretical contribution of this work to legal scholarship is in two areas. First, it will enable traditional policy makers, who seek to modify behavior through regulation, to account for its effect on both deliberative and nondeliberative choice processes. Second, any behavioral attempt to regulate the behavior of people will have to account for the various behavioral trade-offs imposed by normative considerations, for the relative advantages of traditional enforcement mechanisms, and for institutional constraints imposed by the legal culture of different states. This will enable the core ideas of ethical decision making and automatic behavior to be applied to broader societal problems. In the long term, we hope that the combination of behavioral ethics with the behavioral analysis of law will create a new branch of legal scholarship, in which scholars with detailed knowledge of legal doctrines become involved in the theory, and subsequently the practice of mechanism design and behavioral engineering. Behaviorally trained legal scholars will play a prominent role in creating a theory that will ensure greater integrity in the regulation and enforcement of contractual, corporate, and administrative duties.

## [*This part is very preliminary*]

­­­­­­­

## Removed stuff

As part of the need to collect data prior to engaging in a given regulatory tool, we might need to follow the idea developed by Milkman, Chugh & Bazerman (2009) propose the idea that there is a need to match the automaticity intervention and the automaticity of the underlying process that you attempt to modify. This principle should be broadened according to an increased recognition in current legal theory in the justification of personalization according to

## Further research

1. tical overview of the behaviorally informed enforcement strategies  
   1. Ido Erev a famous decision theorist challenges the Becker model of enforcement where the propoblity is most important. I think that the behavioral ethics literature suggests that there is a much greater need related to, self deception, related to gradual change, related to social norms, to people’s inability to do what’s right.
   2. Milkman, Chugh & Bazerman (2009) propose the idea that there is a need to match the automaticity of the intervention and the automaticity of the underlying process that you attempt to modify.
   3. Differences between what an organization can do (as most of the literature is of management scholars)
   4. Better to be feared than loved?
   5. Context sensitivity and the variation in the proportion of good and bad people
   6. Competing models of compliance motivation
      1. Incentives is not as instrumental as one might think.
         1. Legal dollars
         2. Incentive matrix
      2. Expressive function of the law
         1. Is it really distinctive from deterrence?
         2. What do we internalize?
      3. Procedural justice [mandated justice paper]
      4. Can we have it all? Why we can, why we can’t
         1. Competing findings - crowding out motivation
         2. Mandated justice
   7. The promise and peril of money- vs. morality-based legal interventions
   8. Do people choose to violate the law?
   9. Dichotomous vs. linear approaches to legal compliance
2. Dual reasoning and the law: The challenges of changing the behavior of unaware individuals
   1. Automaticity and social policy
   2. Why people obey the law? Revisiting the relevance of compliance motivation research in a world of dual reasoning
   3. Can policy makers modify behavior that is based on automatic reasoning?
   4. Regulating situations vs. regulating people vs. regulating interactions
   5. With great power comes great responsibility ! autonomy responsibility

We have argued that even with the attempts of scholars such as Gneezy et al to create a typology of people’s ethicality, it is very hard to predict ex-ante what would work. This is due to dual reasoning unpredictability and individual differences which are not predictive enough and hence an acoustic serration approach should be pushed forward with recognition that there is some mutual distraction between the different methods. Combination of a situational approach, deterrence and expressive approach were Nudge should be saved only to the few cases where deliberative choice is limited. There is so much to be gained from the fact that people will make a choice – taxonomy could help here as well.



1. Although the ex-post part is far more discussed in law, mainly through the concept of negligence (for a classic discussion see Terry, H. T. (1915). Negligence. *Harvard Law Review*, *29*(1), 40-54. [↑](#footnote-ref-1)
2. [↑](#footnote-ref-2)
3. E.g. Tenbrunsel, A. E., Diekmann, K. A., Wade-Benzoni, K. A., & Bazerman, M. H. (2010). The ethical mirage: A temporal explanation as to why we are not as ethical as we think we are. *Research in Organizational Behavior*, *30*, 153-173. [↑](#footnote-ref-3)
4. In chapter \_\_ which deals with individual variance [↑](#footnote-ref-4)
5. For an illustration of this point see Feldman and Smith; Feldman, Yuval, and Doron Teichman. "Are all contractual obligations created equal." Geo. LJ 100 (2011): 5; Feldman, Yuval, and Shahar Lifshitz. "Behind the Veil of Legal Uncertainty." Law & Contemp. Probs. 74 (2011): 133. [↑](#footnote-ref-5)
6. Ayres, I., & Braithwaite, J. (1995). *Responsive regulation: Transcending the deregulation debate*. Oxford University Press on Demand. [↑](#footnote-ref-6)
7. It make sense to remember at this stage, the earlier works on learning conducted by Skinner, whose approach to learning is such that focused directly (albeit not explicitly) on changing people’s automatic reasoning (think of pablov’s famous findings regard the effect on dogs reaction to noise associated with food. In that regard much of the research on learning tend to [↑](#footnote-ref-7)
8. 1989 JPSP [↑](#footnote-ref-8)
9. See Irene V. Blair, The Malleability of Automatic Stereotypes and Preju dice, 6 Personality & Soc. Psychol. Rev. 242 (2002). Margo J. Monteith et al., Putting the Brakes on Prejudice: On the Development and Operation of Cues for Control, 83 J. Personality & Soc. Psychol. 1029, (2002). [↑](#footnote-ref-11)
10. Her argument is mostly based on the research conducted by E. Ashby Plant & Patricia G. Devine, Responses to Other-Imposed Pro-Black Pressure: Acceptance or Backlash?, 37 J. Experimental Soc. Psychol. 486 (2001). This [↑](#footnote-ref-12)
11. For a review on internalization and socialization processes with regard to wrong doing see Feldman normative failures 2006 attempting to use this literature to understand silicon valley engineers practices of trade secrets’ divulgence. [↑](#footnote-ref-13)
12. See discussion in pages …. [↑](#footnote-ref-14)
13. Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of personality and Social Psychology*, *18*(1), 105. [↑](#footnote-ref-15)
14. Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. [↑](#footnote-ref-16)
15. Marshall, M., & Harrison, S. (2005). It’s about more than money: financial incentives and internal motivation. [↑](#footnote-ref-17)
16. **Ironic effects of anti-prejudice messages: How motivational intervention reduces (but also increases) prejudice. \*Legault, L., \*Gutsell, J. N., & Inzlicht, M. (2011** [↑](#footnote-ref-18)
17. Tenbrunsel, A. E., & Messick, D. M. (1999). Sanctioning systems, decision frames, and cooperation. *Administrative Science Quarterly*, *44*(4), 684-707. [↑](#footnote-ref-19)
18. One caveat is that in many cases external rewards can enhance intrinsic motivation. The interpersonal context in which the extrinsic motivation is introduced, or even the verbal cues attached to the sanctions, can determine how much we intrinsically value the extrinsic reward. For example, a child being reprimanded by a parent, whose opinion the child greatly values, may experience an greater increase in motivation to behave well than if the same reprimand were issued by a teacher with whom the child has little rapport. Nevertheless, the consensus in the literature suggests that in most instances attempts to externally control people’s behavior can have considerable counterproductive results in the long term. (For a review of some of these conflicting effects, see Deci et al. (2001). [↑](#footnote-ref-20)
19. 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, 835–848. [↑](#footnote-ref-21)
20. Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice.

    Journal of Personality and Social Psychology, 75, 811–832. [↑](#footnote-ref-22)
21. But this is not the case with regard to blood donation. [↑](#footnote-ref-23)
22. *Discuss*  Tirrole; Benabu. [↑](#footnote-ref-24)
23. Refer to discussion above. [↑](#footnote-ref-25)
24. It would be noted that when speaking about trade secrets, the focus here is on the narrow definition of trade secrets focusing on the core proprietary knowledge of a given company. Elsewhere, [one of us Orly Lobel, Talent Wants to be Free (2013)] has argued for the importance of information spillover between firms. In that regard we wish to narrow our argument for the core knowledge of firm, rather than the types of information which would be better off shared among companies. See also, YuvaL Feldman, " *The Expressive Function of the Trade Secret Law: Legality, Cost, Intrinsic Motivation and Consensus"* 6(1) Journal of Empirical Legal Studies, 177 (2009). [↑](#footnote-ref-26)
25. This argument is obviously oversimplified and tuning is highly needed here. [↑](#footnote-ref-27)
26. 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. [↑](#footnote-ref-28)
27. Feldman, Y., & Nadler, J. (2006). The law and norms of file sharing. *San Diego L. Rev.*, *43*, 577. [↑](#footnote-ref-29)
28. Connect here to the work on the context of moral identity [↑](#footnote-ref-31)
29. The incentive matrix [↑](#footnote-ref-32)
30. These included items as moral outrage, legitimacy and perceived risk from the misconduct in this factor. [↑](#footnote-ref-33)
31. At the same time, here too an overly broad definition of trade secrets and misappropriation can have detrimental consquences to innovation. Orly Lobel, Talent Wants to be Free (2013); Yuval Feldman, *The Expressive Function of the Trade Secret Law: Legality, Cost, Intrinsic Motivation and Consensus"* 6(1) Journal of Empirical Legal Studies, 177 (2009). [↑](#footnote-ref-34)
32. [↑](#footnote-ref-35)
33. Balcetis, E., & Dunning, D. (2006). See what you want to see: motivational influences on visual perception. *Journal of personality and social psychology*, *91*(4), 612. [↑](#footnote-ref-36)
34. Haisley, E. C., & R. A. Weber. (2010). Self-Serving Interpretations of Ambiguity in Other-Regarding Behavior. *Games and Economic Behavior*, 68(2), 614-625. [↑](#footnote-ref-38)
35. Dana, J., R. A. Weber, & J. Xi Kuang (2007). Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness. *Economic Theory*, 33(1), 67-80. [↑](#footnote-ref-39)
36. Hsee, C. K. (1995). Elastic Justification: How Tempting But Task-Irrelevant Factors Influence Decisions. *Organizational Behavior and Human Decision Processes*, 62(3), 330-337. [↑](#footnote-ref-40)
37. Fienberg, S. E. (Ed.). (2012). *The evolving role of statistical assessments as evidence in the courts*. Springer Science & Business Media. [↑](#footnote-ref-41)
38. For example: good/bad intuition- dissenting views of David G. Rand, Joshua D. Greene & Martin A. Nowak, *Spontaneous giving and calculated greed*, 489(7416) Nature 427-430 (2012); Shaul Shalvi, Ori Eldar & Yoella Bereby-Meyer, *Honesty requires time (and lack of justifications),* 23(10) Psychological Science 1264-1270 (2012). [↑](#footnote-ref-42)