PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO

Guilherme da Franca Couto Fernandes de Almeida

BREAKING RULES

An experimental investigation of the concept of rule

Tese de Doutorado

Thesis presented to the Programa de Pós-graduação em Direito of PUC-Rio in partial fulfillment of the requirements for the degree of Doutor em Direito.

Advisor: Prof. Noel Struchiner
Co-advisor: Prof. Ivar Rodriguez Hannikainen

Rio de Janeiro
December 2020

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO

Guilherme da Franca Couto Fernandes de Almeida

BREAKING RULES

An experimental investigation of the concept of rule

Thesis presented to the Programa de Pós-graduação em Direito of PUC-Rio in partial fulfillment of the requirements for the degree of Doutor em Direito.

Prof. Noel Struchiner
Advisor
Departamento de Direito – PUC-Rio

Prof. Ivar Rodriguez Hannikainen
Co-advisor
Department of Philosophy – Universidad de Granada

Prof. Fábio Carvalho Leite
Departamento de Direito – PUC-Rio

Prof. Fábio Perin Shecaira
Departamento de Direito – UFRJ
Departamento de Filosofia– UFRJ

Prof. Joshua Knobe
Department of Philosophy – Yale University
Program in Cognitive Science – Yale University

Prof. Walter Sinnott-Armstrong
Department of Philosophy – Duke University

Rio de Janeiro, December 4th, 2020

Guilherme da Franca Couto Fernandes de Almeida

Graduated Law School at Universidade Federal do Rio de Janeiro (UFRJ) in 2014. Concluded the master's degree program in Law at PUC-Rio in 2017. Currently, works as the lead researcher of "Supremo em Números" at FGV Direito Rio, where he teaches courses in Legal Data Science and Legal Psychology.

Ficha Catalográfica

Almeida, Guilherme da Franca Couto Fernandes de

Breaking Rules: an experimental investigation of the concept of rule / Guilherme da Franca Couto Fernandes de Almeida ; advisor: Noel Struchiner; co-advisor: Ivar Rodriguez Hannikainen. – Rio de Janeiro: PUC-Rio, Departamento de Direito, 2020.

139 f. : il. color. : 30 cm

Tese (doutorado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Direito.

Inclui referências bibliográficas

1. Direito – Teses. 2. Regras. 3. Interpretação jurídica. 4. Debate Hart-Fuller. 5. Filosofia experimental do direito. 6. Teoria geral do direito. I. Struchiner, Noel. II. Hannikainen, Ivar. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Direito. IV. Título.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

The ideas expressed here are fruit of an intense collaboration with my advisors, Noel Struchiner and Ivar Hannikainen. For their generosity, I owe a debt of gratitude that goes much beyond ordinary advisory relationships.

Way before our collaboration started, Noel already had the idea of running experiments about the concept of rule. The appropriate conceptual framework, the sense of which questions are important, the main implications that we could derive from experimental results, among many other important features of the papers we came to publish alongside Ivar sprung primarily from reflections that he started long before anyone else was involved. It was thus a joy when Noel welcomed me into NERDS, the research group he leads. I was still a student in the master's degree program, but with help from the many talented people at NERDS, I was able to contribute a little bit to this (to my mind) exciting research program. In many ways, this dissertation represents an attempt to systematize the results of this joint project led by Noel. In every step of the way, I benefited immensely from insightful conversations with Noel about philosophy. I probably abused his patience more than most PhD students did. At some points over the last four years, Noel must have woken up to dozens of text messages from me with anguished questions about fine points in general jurisprudence. No matter how late or inappropriate the messages were, Noel always responded with patience and clarity, dispelling my many confusions and pointing me towards the right directions. More often than not, these exchanges, and not solitary reflection, led to improvements in my own arguments and ideas. For all this, I can't overstate the gratitude that I have for Noel. Besides all the things I already described, Noel is also a very dear friend. It is only fair that the first and longest acknowledgment goes to him.

Currently, I teach a course on Legal Data Science in FGV Direito Rio, where I'm also the lead researcher of a group that does empirical research into the Brazilian Supreme Court. I got into FGV because of my ability to do quantitative research, a skill I developed solely because of Ivar. In mid-2016, he taught a seminar to members of the NERDS research group aimed at doing statistical analysis of experimental data. I had absolutely no previous experience with statistics, but that seminar struck a chord with me. Straight away, I started downloading datasets to look for correlations, build linear models, and learn data wrangling in practice. At every step of the way, I tested Ivar's patience with endless text messages at odd hours (which, turns out, is a pattern for me) either asking for help or sharing small

wins. He was always there with helpful tips and encouraging words. I still bug him every time I need to do something sophisticated, and I still learn a lot from his thoughtful answers. In short, I basically owe my career to Ivar. I can't ever be grateful enough for that. Despite my focus in statistical methodology, I can't gloss over the fact that Ivar is first and foremost a world-class philosopher. The arguments, their presentation, the specific examples used to explain each point, as well as the overarching explanations to the data advanced in this dissertation benefited immensely from comments and conversations with Ivar, for whom I'm profoundly thankful.

Perhaps the most euphoric moment during this PhD came when prof. Joshua Knobe accepted to have me as a visiting assistant in research for a few months. Prof. Knobe is probably the most important experimental philosopher in the world. After getting to know him, I was eager to learn that he is also among the kindest and most generous philosophers in the world. Having the opportunity to discuss my ideas with him and with members of his lab has been a joy. Several aspects of the experiments and arguments discussed here were greatly improved because of insightful comments from prof. Knobe, to whom I'm grateful.

In 2013, my life's goal was to become a practicing lawyer. In my spare time, I used to read Gadamer, Zizek, and second-rate Brazilian legal philosophy. With the boldness of someone young and ignorant, I decided to write an ambitious philosophical piece as my farewell to law school. Thankfully, Fábio Shecaira accepted to be my advisor and gently started to undo my certainties by requiring clarity and rigor. After a year's work, I was fully converted to the cause of analytic philosophy. Soon before my graduation, Fábio suggested that I should apply for the masters' degree program at PUC-Rio to study under Noel. Hence, he is to blame for my academic career. I'm not sure if I will ever be able to forgive him for that, but I am no doubt grateful for his encouragement and guidance towards good philosophy.

All of the professors at PUC-Rio have been incredibly helpful during graduate school, and I'm thankful to all of them. I am especially grateful for Fábio Leite, whose discussion of many of the ideas put forth here helped ground our philosophy in firm legal grounds. I have no doubt that, in the future, we will collaborate at the intersection between legal philosophy and Brazilian constitutional law. Moreover, Fábio deserves a special place in this list as a cherished friend.

I would also like to thank Ivar Hartmann and Antônio Maristrello Porto at FGV Direito Rio. Throughout the last three years, FGV Direito Rio allowed me the time and resources needed to complete the philosophical research underlying this dissertation. This included granting me a leave to spend a semester abroad. I am very grateful for their help in making this possible, as well as for insightful discussion of my ideas.

Pedro Chrismann and Danilo dos Santos Almeida recently invited me to start a podcast focused on the intersection between legal philosophy, psychology, and

technology. This was an honor to me. When I joined the masters' degree program at PUC-Rio, Pedro and Danilo were the PhD students whom I looked up to for their careful thinking and deep knowledge of legal philosophy. Today, they are dear friends. Starting "Teoria Impura" afforded me the opportunity to have long weekly conversations with them about very interesting issues. In recent months, they have encouraged me to hijack the podcast to discuss papers and ideas relevant for this dissertation. I have eagerly embraced the possibility. Lucas Miotto Lopes and Andre Bogossian were kind enough to join us in some of these discussions. I am extremely grateful to Pedro and Danilo for this opportunity and for their comments on the ideas developed here.

I would like to thank researchers Alice M. de Almeida, Fernanda Coelho, Felipe Litsek, Alice Régnier, and Ursula Vasconcellos from the NERDS research group for their contributions to the vignettes we used in our experiments.

This dissertation includes the results of a cross-cultural collaboration involving researchers from several countries. For their partnership in this endeavor, I thank Fernando Aguiar, Samantha Bensinger, Piotr Bystranowski, Vilius Dranseika, Bartosz Janik, Markus Kneer, Michael Laakasuo, Alice Liefgreen, Alejandro Rosas Lopez, Maciej Próchniki, Niek Strohmaier, and Kevin Tobia.

Professors Rachel Herdy and Rodrigo Tavares have offered very helpful comments on earlier presentations of some of these ideas. Piotr Bystranowski, Brian Flanagan, Lucas Miotto Lopes, and Kevin Tobia were kind enough to make comments to drafts from which this thesis borrows heavily. I am indebted to their constructive criticism, which helped me avoid argumentative pitfalls.

I am also indebted to undergraduate students who heard half-baked presentations of these ideas in seminars and classes at FGV Direito Rio, UFRJ, and PUC-Rio. Their reactions provided much needed guidance for the best way to present the ideas discussed here.

At FGV Direito Rio, I was fortunate enough to work alongside Fernanda Abreu, Alexandre Almeida, Felipe Araújo, Fernando Correia, Gabriela Gattuli, Fernando Leal, Beatriz Neves, José Luiz Nunes, Thomaz Pereira, Carlos Ragazzo, Kaline Santos, Fernanda Scovino, Isabel Veloso, and Diego Werneck, to mention a few. At several points, their companionship and intellectual support has proved decisive. I'm thankful to all of them for the precious time we spent together.

My parents, Aluizio and Maria Lúcia, as well as my brother, Alexandre, contributed immensely to make this work possible with their unfailing support. From my earliest childhood, they have always encouraged my passions and helped me achieve my goals. Recently, their love provided me with what I needed to face the anxiety of being an academic in a society that decided they no longer want academics, especially in the social sciences. I am profoundly grateful to them.

My friends have been a crucial aid in keeping me sane while working through this PhD. Thus, I would like to thank Tiago Andrade, Vinny Dantas, Carlos Eduardo Lopes, Diego Menescal, and Pedro Nepomuceno for the many happy moments we shared over these last years. I would also like to thank Felipe Amaral, Moysés Ferreira, Pedro Lasneaux, Guilherme Moura, and Lucas Zurli for their companionship in many important moments, most notably during Flamengo matches. Finally, my bandmates Pedro Azevedo and Fillipe Rocha were extremely helpful in letting off steam through music.

Last, but by no means least, I would like to thank my wife, Giovanna. Her support went much beyond what one could expect of any partner. What makes this especially surprising is that she believes knowledge is only valuable when it serves some practical goal. This dissertation is mostly (though not exclusively) aimed at advancing descriptive theses that have no direct practical implications. Hence, I can only make sense of her support for my work as a proof of love. Since I love her deeply, I'm very grateful for this.

Abstract

Almeida, Guilherme da Franca Couto Fernandes; Struchiner, Noel (Advisor); Hannikainen, Ivar (Co-Advisor). **Breaking rules**: an experimental investigation of the concept of rule. Rio de Janeiro, 2020. 139p. Tese de doutorado – Departamento de Direito, Pontifícia Universidade Católica do Rio de Janeiro.

Understanding rules is central to understanding law. However, several issues about the concept of rule remain disputed. One of them was at the center of an important debate in 20th century legal philosophy. Are rules mostly about their texts, as HLA Hart thought? Or are the moral goals pursued by a rule built into its very concept – a position taken by Lon Fuller? Despite many decades of sustained debate, both positions still have proponents and opponents. In this dissertation, I argue that this is partly the result of competing appeals to intuition at the heart of each position. Hart and Fuller evoked thought experiments that elicited conflicting intuitions in each of them. The thought that those reactions to each thought experiment are shared by their target audience is implicit in each author's appeals. In other words, the authors bet that the members of the target audience should have certain beliefs and attitudes. But there is no evidence of whether this is the case. Experimental philosophy provides the tools needed to tackle this lack of evidence. By working out precisely what empirical claims underlie philosophical debate, it is possible to come up with experiments that test the success of those appeals. In this dissertation, I set out to employ those tools to try to break the stalemate between Hart and Fuller over the roles of text and purpose in the concept of rule. After restating the debate as two pairs of conflicting theses that make empirical predictions, I review recent experimental work surveying the intuitions of laypeople and lawyers about rules. These studies involved thousands of participants in several countries and have important implications for the analysis of the concept of rule. They show that text and purpose are prevalent under different circumstances. Moreover, there are important differences between the intuitions of lawyers (who lean textualist) and laypeople (who lean towards purposes). These differences include a cross-cultural divide, with substantial cultural variation among laypeople, but convergence among lawyers. The evidence also suggests entirely new research

questions for general jurisprudence about the precise nature of purposes and interpersonal differences in decision-making style. I consider how each of those findings relate to the broader themes dividing Hart and Fuller (such as the debate regarding the connections between law and morality), as well as the limitations of the existing evidence. Finally, I argue that this model of experimental jurisprudence might be fruitfully applied to other longstanding debates in legal philosophy.

Keywords

Rules; Legal interpretation; Hart-Fuller debate; Experimental philosophy; Experimental jurisprudence; General jurisprudence; The concept of law

Resumo

Almeida, Guilherme da Franca Couto Fernandes; Struchiner, Noel (Advisor); Hannikainen, Ivar (Co-Advisor). **Violando regras**: uma investigação experimental do conceito de regra. Rio de Janeiro, 2020. 139p. Tese de doutorado – Departamento de Direito, Pontifícia Universidade Católica do Rio de Janeiro.

Compreender regras é central para entender o direito. Porém, várias questões envolvendo o conceito de regra permanecem disputadas. Uma delas esteve no centro de um debate importante na filosofia do direito do século XX. Será que regras são fundamentalmente uma função de seus textos, como pensava HLA Hart? Ou será que os objetivos morais perseguidos por regras estão embutidos no próprio conceito – uma posição assumida por Lon Fuller? Apesar de muitas décadas de debate contínuo, as duas posições ainda encontram defensores e detratores. Nessa tese, argumento que isso é em parte o resultado de um conflito entre os apelos à intuições que embasam cada posição. Hart e Fuller invocaram experimentos mentais que elicitaram intuições conflitantes em cada um deles. A ideia de que essas reações aos experimentos mentais seriam compartilhadas por uma população alvo está implícita no apelo que cada autor faz. Em outras palavras, os autores apostam que membros da população alvo têm certas crenças e atitudes. Porém, não há nenhuma evidência a respeito do sucesso dessas apostas. A filosofia experimental oferece as ferramentas necessárias para enfrentar essa falta de evidência. Ao precisar exatamente quais afirmações empíricas subjazem o debate filosófico, é possível conduzir experimentos que testam o sucesso desses apelos. Nessa tese, eu aplico essas ferramentas para tentar resolver o empasse entre Hart e Fuller sobre os papéis de texto e propósito no conceito de regra. Após reformular o debate como dois pares de teses conflitantes que fazem previsões empíricas, eu reviso artigos experimentais recentes que investigam as intuições de juristas e leigos a respeito de regras. Esses estudos envolveram milhares de participantes em diversos países e trazem implicações importantes para a análise do conceito de regra. Eles mostram que texto e propósito são prevalentes sob circunstâncias diferentes. Além disso, eles revelam diferenças importantes entre as intuições de juristas (que são mais textualistas) e leigos (que usam propósitos de forma mais frequente). Essas

diferenças também se manifestam na variação das intuições entre pessoas de culturas diferentes. Leigos de culturas diferentes apresentam diferenças em seus julgamentos a respeito de regras, enquanto juristas do mundo inteiro convergem. Os resultados também sugerem perguntas de pesquisa inteiramente novas para a teoria geral do direito a respeito da natureza precisa dos propósitos e sobre as diferenças interpessoais em estilo de tomada de decisão. Eu avalio a forma como cada um desses achados se relaciona com os temas mais amplos que dividem Hart e Fuller (como o debate a respeito das conexões entre direito e moral), assim como as limitações da evidência existente. Finalmente, argumento que o modelo da filosofia experimental do direito pode ser aplicado de forma frutífera a outros debates longevos em teoria do direito.

Palavras-chave

Regras, Interpretação jurídica; Debate Hart-Fuller; Filosofia experimental; Filosofia experimental do direito; Teoria do direito; O conceito de direito.

Table of contents

List of figures

List of tables

List of Abbreviations

CSM	- Core of Settled Meaning thesis (p. 41)
DNR	- Dual Nature of Rule explanation (p. 69)
ISA	- Indirect Speech Acts explanation (p. 73)
SP	- Strong Purposivist thesis (p. 42)
ST	- Separability Thesis (p. 47)
TP	- Textual Prevalence thesis (p. 44)
WP	- Weak Purposivist thesis (p. 44)

1
Introduction

Imagine that Congress creates a rule with text to the effect that: "A stable union between a man and a woman, evidenced by their public engagement in continuous and durable companionship with aims to constitute a family, is recognized by the state as a family and protected as such"[1]. Under the terms of such a relationship[2], companions inherit from each other and might owe each other financial obligations in case of a break up. One reason to adopt such a rule is to protect people engaged in enduring romantic relationships from unfair outcomes that might arise from their end. This is especially important when one of the partners relies on the other for financial support. For instance: it is unjust that someone who decided to follow less lucrative paths to focus on family life with the willing consent of his or her companion should have to live in underprivileged conditions after the relationship ends.

The rule passed by Congress sometimes (hopefully often) accomplishes this goal, but there are times when text and purpose will come apart. Consider the following hypothetical case. John and George were a happy couple living in Rio de Janeiro. George was a wealthy lawyer and John forwent some lucrative job opportunities outside Rio for the sake of their relationship. The fact that he didn't earn much at his job never bothered John too much. George was wealthy enough for both of them; nurturing a healthy and close relationship was something the couple deemed more important than further financial success. They were living

[1] Loose translation of the Brazilian Civil Code, art. 1723.
[2] A correlate idea is present in common-law countries under the name "common-law marriage".

together for 10 years when George tragically died in a car accident. As he was only 40, he had made no arrangements for his death.

George's parents were very conservative. They were never able to get over the fact that their son was gay. After George's death, they consulted a lawyer, who pointed out that, by the law's text, only a man and a woman could be in a "stable union". As such, they - and not John - were the legal heirs to George's wealth, including the apartment where John currently lived. Moved partly by their dislike of John, they filed suit seeking to inherit their late son's estate.

When John was notified of the lawsuit, he became livid. Upon consultation, his lawyer pointed out that even though George's parents were right about the rule's text, they were very much in the wrong once we accounted for the purposes behind it. If the rule's goal was to protect companions from the unfortunate financial fates they could face after the death of a loved one, surely John – and not George's parents – was the rightful heir.

When examining John vs. George's parents, Richard, the judge charged with the case, will need to determine what shall be done with George's estate. In doing so, he will have to consider what the relevant rule has to say. The law requires him to do so. This is not to say that rule-mandated results can't be overridden: after all, the law might also expect Richard to decide whether an exception should be made, whether the rule is compatible with hierarchically superior norms, and so on. But all of these potentially overriding factors already presuppose a certain rule-mandated result. If the rule doesn't apply, it makes no sense to speak of granting exceptions to it or of ruling it unconstitutional. When pondering about rule application after hearing the arguments put forth by the parties, Richard might become genuinely puzzled. Is a rule identical with its text – as argued by George's parents -, its purpose – as argued by John -, or some combination of both? If he is philosophically curious, he might start investigating the literature that discusses the nature of legal rules, where the Hart-Fuller debate looms large.

By studying the Hart-Fuller debate, Richard will learn many things. Chief among them, given the issue at hand, he will be exposed to many valuable arguments about how text and purpose interact. But he will not find an univocal

answer to the question of whether the rule was violated in situations such as John vs. George's parents. As we will see, Hart would say that the rule was violated, but Fuller would likely disagree. This stalemate persisted through several rounds of debate between the two philosophers and their torch bearers. Why couldn't outstanding philosophers reach a consensus about such an important issue? I believe that systematic data on lawyers' and ordinary citizens' intuitions about rules might help us answer this question. Where philosophical debate is sometimes hindered by implicit and untested assumptions, experimental philosophy might come to the rescue by making these presuppositions explicit and clarifying which of them hold true. More importantly yet, understanding the empirical underpinnings of the Hart-Fuller debate might help achieve a better understanding of the concept of rule. Thus, this dissertation uses the Hart-Fuller debate as the starting point into an experimental investigation of the concept of rule.

But before turning to this project, I would like to note that investigating this specific issue in the concept of rule is a worthwhile project not only for its theoretical value, but also because of its practical importance. Richard's hypothetical dilemma is meant to provide a clean illustration of a very real issue facing real judges every day. In fact, Richard's conundrum was loosely inspired by an issue decided by the Brazilian Supreme Court in 2011[3] where text and purpose recommended different results. On that occasion, the court sided with purpose[4], but the polemics that followed show that the question was far from trivial, and many of those polemics revolved around the text vs. purpose debate (see Nigro, 2013).

Moreover, the interaction between text and purpose within the concept of rule has practical implications much beyond law. Parents, club owners, crime lords, and private administrators all deal with rules on a daily basis. For each of them, clashes between text and purpose might also be consequential, introducing decision-making challenges similar to those that arise in a legal context. The experiments I will review include not only legal (e.g., a rule establishing a speed limit on a highway), but also non-legal rules (e.g., a rule barring people from entering an apartment with their shoes on). So far, the data haven't shown any

[3] Joint judgment of the cases ADPF 132 and ADI 4277.
[4] Even if purpose only played an explicit and central role in the vote of Justice Celso de Mello.

systematic difference between intuitions about these two kinds of rules. As a result, despite my almost exclusive focus on jurisprudence, many of the conclusions and arguments developed here might be useful outside the legal domain.

Finally, accurately describing the ordinary concept of rules provides important guidance for normative questions troubling those who make rules. When drafting rules, should we spend a long time laboring over the exact wording of the rule's text, or should we work towards clearly communicating our intended purpose? This hinges on how people are likely to interpret the demands of the rules we create. So, studying intuitions about rules might also help us make rules that are more effective.

1.1
A brief history of the Hart-Fuller debate and why it matters

The debate between Herbert Hart and Lon Fuller spanned over a decade and is comprised of several different pieces by each author. The best summary of the debate's history lies at its very end, when Fuller used the colorful allegory of boxing rounds to describe his exchanges with Hart:

> It began when Professor Hart published the Holmes Lecture delivered at the Harvard Law School in April 1957. In that lecture he undertook to defend legal positivism against criticisms made by myself and others. The first attempt at counterthrust was my critical commentary on this lecture. Round three was marked by the publication of Hart's *The Concept of Law*; round four occurred when the first edition of the present work [*The Morality of Law*] was published; round five took place when Hart published his review (Fuller, 1969, p. 188).

The excerpt serves as an introduction to Fuller's "A reply to critics", published as the final chapter of a revised version of *The Morality of Law*. The piece marks the last time either author dedicated themselves at length to each other's positions and criticisms[5], thus marking what would be the sixth and final round of the dispute. Along the way, the arguments and central points of contention evolved, but one of the main issues was already set on the very first round of the debate: what are the roles played by text and purpose in the concept of rule?

[5] In the introduction to *Essays in Jurisprudence and Philosophy*, Hart (1983) dedicates a few sentences to his debates with Fuller, but only *en passant*. One of those sentences concedes too much to Fuller, as I will discuss in chapter 2.

During the lecture that initiated the debate (later published in the 1958 edition of the Harvard Law Review), Hart asked us to consider the example of a rule prohibiting vehicles in a public park. Even though we might disagree about whether "bicycles, roller skates, [and] toy automobiles" count as vehicles, a regular car surely does. Moreover, there is no need to resort to moral judgment in order to see that cars fall within the scope of the no-vehicles rule: the rule's text is sufficient to determine that the entrance of cars violates the rule.

According to Lacey (2010), during Hart's presentation, Lon Fuller "paced up and down at the back of the room 'like a hungry lion' and later demanded a right to reply" (pp. 1-2). He had several problems with Hart's positivism, and one of them related to the interpretation of rules. For him, morality played a role in interpreting the no-vehicles rule even in clear cases:

> What would Professor Hart say if some local patriots wanted to mount on a pedestal in the park a truck used in World War II, while other citizens, regarding the proposed memorial as an eyesore, support their stand by the "no vehicle" rule? Does this truck, in perfect working order, fall within the core or the penumbra (Fuller, 1958, p. 663)?[6]

Schauer argues that, "In offering this example, Fuller meant to insist that it was *never* possible to determine whether a rule applied without understanding the purpose that the rule was supposed to serve" (2008, p. 1111). In other words, for Fuller, it is impossible to know whether or not the rule applies without looking for its purpose. Thus, contrary to what Hart believed, Fuller argued that analysis of text is *never* sufficient for ascertaining rule violation. This form of purposivism entailed conclusions to which Hart could not concede[7]: if people necessarily resort to morality in order to ascertain the content of law, positivism – a position famously associated with Hart – fails[8].

This and many other arguments discussing the original pair of examples relies partly on the fact that the "no vehicles in the park" rule is specific (as opposed to vague), has a canonical textual formulation, and applies in a roughly all-or-

[6] Incidentally, some people do deny that this truck is a vehicle, pointing out that "If it cannot move, it might be said, it is not a vehicle" (Schauer, 2008, p. 1116).
[7] But see Bix: "Hart would accept Fuller's position as the starting-point for judicial interpretation, but he would disagree that such resources are sufficient to avoid the need for judicial discretion" (2003, p. 29).
[8] The nature and implications of the disagreements between Hart and Fuller about the concept of rule are developed much further in chapter 2.

nothing fashion (according to the rule, either something is or isn't a vehicle – there is no intermediate option)[9]. Specificity is important because both Hart and Fuller want to argue about things that happen when the rule's text unambiguously covers certain cases. This is relatively easy to do when we are dealing with rules enunciated in clear text, such as the "no vehicles in the park" rule. Where a rule's text is vague, or ambiguous, on the other hand, it is much harder to probe where the influence of text ends and where purpose takes over. Canonicity is important for similar reasons. If "no vehicles in the park" was an unwritten rule, disputes might arise as to the wording of the rule's text in such a way as to threaten the possibility of discussing the sufficiency or insufficiency of text for rule violation. Finally, the fact that the "no vehicles in the park" rule applies in an all-or-nothing fashion matters because it allows Hart and Fuller to frame their examples in terms of simple compliance or non-compliance. Where rules invite considerations of weight, the relationship between text and purpose might turn out to be much more complex.

All of this matters to clarify the scope of the Hart-Fuller debate and of this dissertation. Although rules vary wildly across these three dimensions – contrast a rule establishing a speed limit with a highly abstract implicit constitutional principle – my attention will focus on specific, canonical and all-or-nothing rules. Moreover, when Hart and Fuller discuss the role of text in legal interpretation, they do not mean it to include all textual sources within a legal system, or even in a given piece of legislation. As illustrated by the "no vehicles in the park" rule, they mean to debate whether the text of the most locally applicable rule is sufficient to determine that this specific canonical rule was violated[10]. Thus, this dissertation provides an experimental investigation of the concept of rule in this narrow sense.

Ever since Hart and Fuller began their exchange, there was ample academic interest regarding this specific issue in their broader divide, starting with their own

[9] For an analysis on how those three characteristics might be more important to classify legal prescriptions than the rule-standard dichotomy, see Schauer (1996). In a slight departure from Schauer, I am going to use "rule" in what follows to denote something that lies roughly at one end of the spectrum in all three dimensions. In other words, I'm going to use the term "rules" to refer to prescriptions that are specific, canonical and fairly weighty.

[10] Moreover, violation of the rule might not even entail violation of the *law*. For instance, a judge might find that, even though the rule was violated, it is incompatible with the constitution, or with other legal requirements. In that case, the fact that the rule was violated will become legally irrelevant.

writings in later rounds of the debate[11]. A brief and certainly non-exhaustive list of works by other philosophers that discuss the "no vehicles in the park" examples and their implications to legal interpretation and general jurisprudence includes Barak (2005), Bix (1991, 2003), Fish (2006), Flanagan (2010), Hurd (2015), Levin (2012), Marmor (2005, chapter 7), Schlag (1999), Shecaira (2015), Slocum (2015, chapter 5), Soames (2012), Tobia (forthcoming), and Vega Gomez (2014). What's more, the Hart-Fuller debate has set the scene for the literature that followed it: Schauer's (1991) influential take on the concept of rule uses text and purpose as the analytical components needed to define rules and rule-based decision-making; works on legal interpretation (such as Shecaira, 2009; Barak, 2005) routinely emphasize these two components as especially relevant.

But who was right after all? Most contemporaneous analytic philosophers in the Anglophone tradition seem to take Hart's side[12]. But even among this very restricted demographic, many still defend purposivist views. For instance, it is possible to argue that Dworkin's theory of legal interpretation is a purposive theory following in the footsteps of Fuller[13].

At this point, one objection to this dissertation's project might become salient: if sixty years of debate among the brightest legal philosophers in the world weren't enough to put the issue to rest, adding one more PhD dissertation mustering philosophical arguments for either side could hardly move the needle. This might

[11] The other rounds of the Hart-Fuller are comprised of Hart's *The Concept of Law* (1994, originally published in 1961), Fuller's *The Morality of Law* (1969, originally published in 1964), Hart's 1964 review of *The Morality of Law* (reprinted in Hart, 1983, chp. 16), and Fuller's response to critics in the revised edition of *The Morality of Law* (1969, chp. V).

[12] Lacey (2010, p. 4): "a fair-minded observer of the jurisprudential scene would have to conclude that Hart, as it were, won the war". Even those who defend Fuller tend to acknowledge this point, see Burg (2014), Rundle (2012).

[13] See Schauer (2008, pp. 1130-1131): "One way of understanding Fuller, and possibly theorists such as Ronald Dworkin and Michael Moore as well, is believing that the good judge is one who sets aside the plain language of the most directly applicable legal rule in the service of purpose, or of reasonableness, or of making the law the best it can be, or of integrity, or simply of doing the right thing". Shecaira (2014, p. 26) offers a similar reading: "[Dworkin] believes that the interpretation of legal sources inevitably involves moral deliberation (and that is where he departs from positivism)". Burg (2014, p. 741) comments how "there are many commonalities between Fuller and Dworkin – ones that Dworkin strangely enough has never bothered to discuss". Bustamante (2020) further specifies the claim that moral purposes are central for Dworkin's theory of interpretation stating that: "[…] what guides the interpretative attitude of a Dworkinian interpreter is a *moral purpose*, a *value* that is capable of providing a justification of the practice being interpreted" (p. 123). Rundle (2012) dedicates chapter 7 of her book to an evaluation of the similarities between the projects of Fuller and Dworkin.

very well be true. My hope of proving this form of skepticism wrong hinges on a methodological point: by looking beyond the methods of armchair philosophy towards an empirical perspective, we might be able to make genuine progress (or, at the very least, we might fail in an original way). Just as experiments have proven illuminating with regards to the concepts of intentional action (Kneer & Burgeois-Gironde, 2017; Tobia, 2020a), reasonableness (Tobia, 2018), consent (Sommers, 2020), ordinary meaning (Tobia, 2020b), causation (Macleod, 2019; Knobe & Shapiro, forthcoming), and law itself (Donelson & Hannikainen, 2020) – all very contentious and relevant for legal philosophy -, they can also help with the concept of rule[14]. And insofar as the concept of rule is at the core of one of the "persistent questions" (Hart, 1994) that arise in the quest for the concept of law, this experimental strategy might help advance the more ambitious project of general jurisprudence – as foreshadowed by the implications underlying Fuller's rebuttal to Hart.

Hence, this dissertation's title "Breaking rules" refers both to the topic (the concept of rule and rule violation judgments) and to the method of the investigation, since it represents a break with the way contemporary legal philosophers interested in general jurisprudence approach their research questions. But how exactly can this methodological shift help us make better philosophy?

1.2
Experimental philosophy as a path forward[15]

One strategy frequently employed by philosophers, legal and otherwise, is what some have called the "appeal to intuition":

> A philosopher describes a situation, sometimes real but more often imaginary, and asks whether some of the people or objects or events in the situation described have some philosophically interesting property or relation [...]. When things go well, both the philosopher and her audience will agree on an answer, with little or no conscious reflection, and they will take the answer to be *obvious*. The answer will then be used as evidence for or against some philosophical thesis (Stich & Tobia, 2016. p. 6).

[14] For an overview of the young field of experimental jurisprudence, see Tobia, 2020c.
[15] The arguments in this section closely follow the structure and examples used in Almeida, Struchiner & Hannikainen (under review).

This strategy is prominently at play in the Hart-Fuller debate: Hart [Fuller] describes an imaginary situation (a regular car enters the park [a group mounts a functional truck as a monument in the park] under the "no vehicles in the park" rule) and expects his audience to share his intuition that the car is [the truck isn't] in violation of the rule. These intuitions are then taken to express something fundamental about the concept of law itself (law can [can't] be a purely factual matter). As a result, Hart's concept of law emphasizes the separation of law and morality, while Fuller's theory focuses on the moral purposiveness of law.

When we cast the debate in these terms, one thing stands out: Hart and Fuller appeal to different and incompatible intuitions[16]! Hart himself sometimes acknowledged this incompatibility, stating at a given point that he was "haunted by the fear that our starting-points and interests in jurisprudence are so different that [Fuller] and I are fated never to understand each other's work" (1983, p. 343). From this standpoint, it is no wonder that so many rounds of debate ended in stalemate. After all, each philosopher's theory was tailored to give center stage to his own set of intuitions, while relegating the other set to a supporting role. However, if this is the best explanation for the stalemate, there might be a way out. We might just replace the assumption that our audience agrees with our intuitions with experimental tests of this hypothesis. This is precisely what experimental philosophers in the "positive program" do:

> [Experimental philosophers under the positive program] can avoid some of the idiosyncrasies, biases, and performance errors that are likely to confront philosophers who attend only to their own intuitions and the intuitions of a few professional colleagues who read the same journals and who may have prior commitments to theories about the concepts under analysis. By collecting the intuitions of a substantial number of nonphilosophers, Knobe maintains, we may discover important facts about ordinary concepts that have gone unnoticed by philosophers using more traditional methods of conceptual analysis (Stitch & Tobia, 2016, p. 10).

Whereas armchair philosophers' bets about human intuitions might be wrong, experimental philosophers avoid losses by not entering losing wagers. Inasmuch as the appeals to intuition made by traditional philosophers are sincere, this should be a welcome change. It would help keep our theories true by weeding

[16] This is true even if the original pair of examples isn't able to show the incompatibility, an issue I will discuss on chapter 2.

out intuitions that turn out not to be shared by a large majority. If an armchair philosopher makes an appeal to intuition and doubles down on his position after it is shown that his intuitions aren't widely shared, his appeal is only an insincere, rhetorical flourish. So, if we could show Hart (or Fuller) that his intuitions are not shared by the population at large – or that they are shared by only some specific subset of the population under some specific circumstances -, he would presumably recant and reconsider his theory working from another set of facts.

Even though some might be skeptical of the usefulness of experimental methods to advance philosophy in general[17] - because, for instance, the intuitions that surveys are capable of collecting are irrelevant for understanding concepts -, I believe that legal philosophers are much less likely to show skepticism[18]. After all, several of the most prominent legal philosophers featured in contemporary debates agree that the mental states of people shape law: "Each of the major conceptual theories of law – positivism, natural law theory, and Dworkin's interpretativism – begin from the assumption that law is, in part, manufactured by contingent social practices […]" (Himma, 2015 , p. 86). These practices amount to things like having certain behaviors (e.g., to obey the legal rules), beliefs (e.g., this specific set of rules count as legal, while this other does not) and attitudes (e.g., to accept legal rules as guides to evaluate behavior), all of which can be measured using empirical methods. Hart and Fuller are no exceptions in that both recognize the importance of those mental states to legal philosophy.

Let us begin with Hart. Hart's doctrine of the rule of recognition means, in short, that law is demarcated by certain social facts about the beliefs and attitudes of a system's citizens and officials. His crudest example - entertained only in passing - regards an absolute monarch named Rex (p. 114). Imagine that all of Rex's officials are willing to conform their behavior to each and every rule enacted by Rex and are willing to punish deviations from those rules. It is this attitude by Rex's officials, coupled with general obedience by citizens at large, that makes it so that Rex's rules are law. If suddenly, for some reason, every official and citizen shifted their attitudes of acceptance and obedience to the rules enacted by Brutus instead,

[17] See, for instance, Williamson (2011; 2019).
[18] But see Schauer's (2020) depiction of Raz's metaphysical methodology.

then Brutus's rules would become the legal rules in that system, while any further rules from Rex would lack that status[19]. Thus, what counts as law is purely a matter of the behaviors, beliefs, and attitudes of a given group of people.

This is enough to show that Hartians shouldn't be completely averse to the prospect of experimental jurisprudence. Insofar as they value the study of the rule of recognition, they should welcome systematic studies investigating the behaviors, beliefs, and attitudes that identify it. But why should they welcome experiments specifically about the concept of rule? Because the practice of recognition is mostly about rules: the debate is whether ordinary citizens comply with the legal rules, what set of rules count as legal, which behaviors should be reprimanded, and so on[20]. Reflecting the centrality of the concept of rule for general jurisprudence, Hart commented that: "[…] dissatisfaction, confusion, and uncertainty concerning [the] seemingly unproblematic notion [of rule] underlies much of the perplexity about the nature of law. What *are* rules? What does it mean to say that a rule *exists*?" (1994, p. 8). Hence, understanding intuitions about rules might be necessary to comprehend the concept of law.

What about Fuller? Did he disagree with Hart on this point? I believe he did not. His disagreement with Hart was not about whether facts about people's behaviors, beliefs, and attitudes influenced the content and structure of legal systems. He agreed that they did. His problem was with the best way to characterize the content of these mental states. For him, the beliefs and attitudes at play were moral in nature. Commenting on an early version of Hart's doctrine of the rule of recognition, Fuller wrote:

> When I reached this point in his essay [where Hart notes that the foundation of a legal system rests on certain "fundamental accepted rules specifying the essential lawmaking procedures"], I felt certain that Professor Hart was about to acknowledge an important qualification on his thesis. [...] The question must now be raised, therefore, as to the nature of these fundamental rules that furnish the framework within which the making of law takes place. On the one hand, they seem to be rules, not of law, but of morality. They derive their efficacy from a

[19] Hart's rule of recognition is much more complex than this. What precise roles must be played by officials and citizens in accepting it, what is the rule's relationship to other legal rules, whether or not it can be used to explain legal authority, among many other questions, are still hotly debated. This example is used only to show the (to my mind) modest claim that Hartian positivism holds that law is a function of the mental states of people.

[20] For a deeper debate about the relationship between rules and the rule of recognition, see 3.5.1, infra.

> general acceptance, which in turn rests ultimately on a perception that they are right and necessary. They can hardly be said to be law in the sense of an authoritative pronouncement, since their function is to state when a pronouncement is authoritative. [...] Here, then, we must confess there is something that can be called a "merger" of law and morality, and to which the term "intersection" is scarcely appropriate. Instead of pursuing some such course of thought, to my surprise I found Professor Hart leaving completely untouched the nature of the fundamental rules that make law itself possible [...] (Fuller, 1958, p. 639).

Fuller might have objected to the social aspects of the rule of recognition, but he never did so. He simply objected to the positivist thesis that the rule of recognition isn't necessarily moral: social facts, although important, are not all there is when it comes to law; moral evaluation is also a necessary element in legal recognition.

In fact, if the choice is between a philosophical outlook that gives center stage to ordinary intuitions versus a more metaphysically ambitious enterprise that might do away with these foundations, Fuller unambiguously favors the former. He explicitly rejected the view according to which jurisprudence should be aimed at grasping definitions and necessary truths. In his "A reply to critics", he complains about this tendency of legal philosophy: "Perhaps in time legal philosophers will cease to be preoccupied with building 'conceptual models' to represent legal phenomena, will give up their endless debates about definitions, and will turn instead to an analysis of the social processes that constitute the reality of law" (1969, p. 242)[21]. A focus on the "social processes" presumably should be sensitive to the beliefs and attitudes that people in society actually have.

Thus, for the purposes of the discussion pursued in this dissertation, there seems to be great convergence as to the relevance of the beliefs and attitudes held by people about the law. Inasmuch as experimental methods are useful in uncovering what those beliefs and attitudes are, experimental jurisprudence should be welcomed by most legal philosophers, including Hartians and Fullerians.

Given this common ground, a new question becomes crucial: whose beliefs and attitudes matter? If something like my rough summary of the Hartian rule of recognition obtains, there is some group of people in each legal system whose

[21] This excerpt is repeatedly quoted in Rundle (2012) to illustrate this aspect of Fuller's critique of his interlocutors in jurisprudence.

acceptance[22] of certain rules makes it so that law exists and is structured in a specific way. Which group is it? Adler (2006) coined the expression "recognitional community" to refer to this question. While there is agreement around the idea that beliefs and attitudes shape law, there is deep disagreement about the appropriate recognitional community:

> Is the recognitional community for the U.S. legal system: (1) all living persons; (2) all living persons physically located within what is generally acknowledged to be the territory of the U.S.; (3) all living persons generally acknowledged to be U.S. citizens; (4) all living persons generally acknowledged to be officials in the U.S. legal system; or (5) all living persons generally acknowledged to be judges in the U.S. legal system? Hart in *The Concept of Law* answers, all living U.S. officials; Raz and Kutz answer, all living U.S. judges. Deep popular constitutionalism answers, all living U.S. citizens (p. 726).

What are the positions of Hart and Fuller on that matter? As the quote suggests, the orthodox reading of Hart assigns him the position that only officials belong to the recognitional community. Here is the relevant portion of "The Concept of Law":

> There are therefore two minimum conditions necessary and sufficient for the existence of a legal system. On the one hand, those rules of behavior which are valid according to the system's ultimate criteria of validity must be generally obeyed, and, on the other hand, its rules of recognition specifying the criteria of legal validity and its rules of change and adjudication must be effectively accepted as common public standards of official behaviour by its officials. The first condition is the only one which private citizens *need* satisfy: they may obey each for 'his part only' and from any motive whatever; though in a healthy society they will in fact often accept these rules as common standards of behaviour and acknowledge an obligation to obey them, or even trace this obligation to a more general obligation to respect the constitution. The second condition must also be satisfied by the officials of the system. (1994, pp. 116-117).

At first glance, then, a Hartian empirical investigation should begin and end with the intuitions of officials. However, if we take Hart to be a philosopher concerned not only with necessary truths[23], but also with what goes on in typical and "healthy" societies, we should expand our investigation to include laypeople. Moreover, Hart was by no means steady in his position about the appropriate recognitional community. Earlier, in his 1958 essay, while sketching out a preliminary version of the rule of recognition, he remarked that "[These

[22] Acceptance here refers to a specific set of attitudes specified in Hartian theory. See Shapiro (2006).
[23] More on this on chapter 2.

fundamental accepted rules specifying what the legislature must do to legislate] lie at the root of a legal system, and what is most missing in the utilitarian scheme is an analysis of what it is for a *social group and its officials* to accept such rules" (p. 603, emphasis added). Later, in his postscript to "The Concept of Law", he stated that the rule of recognition "is in effect a form of *judicial* customary rule existing only if it is accepted and practiced in the law-identifying and law-applying operations of the *courts*" (1994, p. 256, emphasis added). Thus, at different points in time, Hart argued that the recognitional community is comprised of citizens and officials alike, only of officials, and only of judges.

Hence, looking at Hart alone, we would already have reason to investigate all kinds of different populations (laypeople, officials, and judges). But the need to survey all kinds of intuitions and discriminate between them is greatly amplified by Fuller's perspective. According to Rundle, Fuller's objection to positivism "[…] is not that it looks to and takes its cues from the source of lawgiving power. The problem, rather, is how positivism privileges that perspective *exclusively*, with no meaningful regard for the way that law's distinctive mode of interacting with *the legal subject*, as an agent, provides the basis for distinguishing law from other modes of rule" (2012, p. 117). Dealing specifically with interpretation, Rundle interprets Fuller as maintaining that "interpretation is, above all, not about individual words[24], but about taking on responsibility for maintaining laws that are intelligible to *their subjects*' situation" (p. 170). Several other aspects of Fuller's work suggest this reading. For instance, his requirement that laws should be promulgated hinges on the *citizen's* right to have access to the content of laws[25].

Thus, including lawyers and laypeople in our experimental studies and discriminating between them in analysis is of paramount importance. Doing so will either lead to the conclusion that intuitions across all groups converge, rendering the debate around the recognitional community futile, or reveal systematic differences between groups. In case systematic differences exist, knowing exactly where they diverge might help us move forward. It might prove useful, for instance,

[24] Schauer, 2008, convincingly argued that Fuller's claims that Hart's position hinges on the interpretation of "individual words" is misguided.
[25] "Even if only one man in a hundred takes the pains to inform himself […] this is enough to justify the trouble taken to make the laws generally available" (Fuller, 1969, p. 51).

in understanding enduring debates in legal philosophy. Perhaps groups of thinkers that have divergent communities in mind end up with different concepts of law precisely because of the different beliefs and attitudes that these communities have[26]. Differences between communities might also have practical repercussions. If some groups are textualist, while other are purposivist, this should be accounted for during the decision-making process of judges, and the rule-drafting process of lawmakers.

With all of this in mind, I took an ecumenical approach to the recognitional community question. Wherever possible, I surveyed data from studies involving both people who are lay about law, and legally trained subjects. Unfortunately, no studies conducted so far about the concept of rule focused on legal officials properly[27], but, presumably, studying lawyers – a subset of officialdom – might give us some insight into their psychology[28]. This approach proved fruitful, as we uncovered important differences between the intuitions of laypeople and lawyers about the concept of rule.

There is still another sense in which the "whose intuitions matter" question is important. An ultimate rule of recognition is a practice that defines the existence of a *particular* legal system. As such, it can be used to distinguish between different legal systems: the legal systems of Brazil and Oman are different because they rest on different sets of ultimate practices[29]. As someone engaged in general jurisprudence, however, Hart was much more interested in the similarities between the vastly different legal systems of Brazil and Oman than in their differentiating traits. Put another way, the question animating Hart (and also Fuller) was: what makes these two systems containing very different rules two instances of law? Both Hart and Fuller would answer this question by pointing to certain shared features of the mental states involved in people's interactions with legal rules no matter where. Hart's account would emphasize the fact that people don't necessarily make

[26] Adler (2006) himself makes some suggestions in that direction.
[27] Hart's account of officials included not only judges, but also other officers responsible for applying the law, such as legislators and police officers (see Adler, 2006). However, Hart is imprecise in his references to "officials" (see Hubbard, 2011, pp. 94-95).
[28] For a slightly more extended argument that this is the case, see 3.5.2, infra.
[29] See the discussion in Hart (1983), Essay 15.

a moral judgment to identify legal rules, while Fuller would insist that legal interpretation always involve appeals to moral purposes.

Note that these two answers assume that the mental states that go into rule-based reasoning are similar everywhere. But are they? Or do different cultures hold such different concepts of rule as to make general jurisprudence an impossibility? These are questions that necessitate cross-cultural investigation. In order to make experimental general jurisprudence, it isn't sufficient to investigate how people deal with rules in one specific jurisdiction: it is necessary to study how they think across the globe. If people share one single concept of rule, then it might be possible to advance one single theory that accounts for law everywhere, as general jurisprudents such as Hart and Fuller wished to do. If, however, they have highly diverse intuitions about what constitutes a rule, we might have to abandon this project altogether[30].

The cross-cultural project and the recognitional community question are distinct enterprises. Maybe laypeople all over the world share the same intuitions about rules, but lawyers' intuitions vary by culture. Maybe the inverse is true: culture might heavily influence laypeople while keeping lawyers untouched. Another possibility is that lawyers and laypeople always agree within each country, but there is variation across countries. And so on. How should we assess the viability of general jurisprudence in this context? One suggestion is that general jurisprudence is viable if at least one recognitional community has cross-culturally robust intuitions about legally important concepts, chief among them, that of rule. Hence, combining an ecumenical approach to the recognitional community with a cross-cultural scope seems to be the most promising path for experimentalists concerned with general jurisprudence.

1.3
The plan

I hope to have convinced the reader that experimental jurisprudence is worth a shot. If I succeeded, there is still much work to be done. Maybe the first and foremost question that needs to be answered is the following: how can empirical

[30] For more on this, see chapter 4.

evidence help adjudicate between the *conceptual* claims at play in the Hart-Fuller debate? Hart and Fuller weren't experimentalists. They haven't explicitly suggested ways in which experimental data might prove them right or wrong. To show that their debate can be advanced through experimental means involve spelling out the predictions that could falsify each philosophical position.

The first chapter tries to rise to this challenge by working out the empirical claims underlying the Hart-Fuller debate and their broader implications for jurisprudence. More precisely, I will argue that the conceptual positions advanced by the authors can be expressed as four empirical theses organized in two pairs of opposing claims. Furthermore, one of those pairs of theses crosses over a popular way of characterizing legal positivism. Finally, I will argue that Schauer's (1991) work on rules offers tools that can be used to test these four theses.

In the second chapter, I present the evidence collected to date, both by my supervisors and myself (in Struchiner, Hannikainen & Almeida, 2020; Almeida, Struchiner & Hannikainen, under review), as well as by others (such as Turri & Blow, 2015; Turri, 2019; Garcia, Chen & Gordon, 2014). The data show that both appeals to intuition have some meat to them: both text and purpose influence people's judgments about rules. However, circumstances about the decision-making context influence which one prevails; under some conditions, people are textualists, while under other conditions, they are purposivists. Moreover, there are important differences between the intuitions of lawyers and lay people. As foreshadowed by the discussion of recognitional communities in the introduction, these differences might turn out to explain at least some of the differences between Hartian and Fullerian scholars.

The third chapter discusses the possibility of general jurisprudence. Hart and Fuller were concerned with mapping the concept of law in all human societies. Recently, some have pushed back against the possibility of such an enterprise (Schauer, 2015; Tamanaha, 2001), arguing that the wide-ranging cultural variation among legal systems makes it hard to formulate one single concept of law that is also interesting. Applied to the discussion regarding the concept of rule, this is an empirical question: do people all around the world share a single concept of rule, or is there significant cultural variation? How does this data interact with the

question of the relevant recognitional community? Here, I present preliminary results of an ongoing work (Almeida et al., in prep) surveying the intuitions of people in a number of different countries around the globe. The results suggest that there is ample cultural variability among laypeople, but not among lawyers.

The fourth chapter explores two venues of philosophical inquiry that spring out from the existing data. First, I look more closely at purposes to question the role they play in the concept of rule. In legal philosophy, as well as in cognitive science, there are multiple concepts of purposes. Some argue that purposes are one and the same with the original intentions of whoever created a rule or a physical object. Others maintain that purposes are necessarily moral. Finally, some people argue that purposes are identified by current practices surrounding a rule or an object. Which one of these versions of the concept of purpose is relevant to rule violation judgments? How do people apply purposes when they are left implicit? These are important questions that are suggested by empirical investigation, but that are not salient in the traditional, armchair debate about rules. Second, I consider whether the results surveyed in chapters two and three are best understood as caused by different traits – some people are Hartian throughout, while other are Fullerian – or states – sometimes, we lean Hartian, other times we lean Fullerian. If the differences are caused by traits, the best description of law will always depend on features about the population of each jurisdiction. Considering the possible decision-making profiles that arise from the trait hypothesis reveal some stances that have been overlooked by traditional philosophical debates, but that are nonetheless normatively interesting. If, on the other hand, the data captures different states of mind, people might be nudged one way or another by features of each situation and the best description of law should take these contextual variations into account.

In the conclusion, I will revisit the argument for the use of experimental methods in general jurisprudence. Not only can empirical methodology help move forward existing debates (chapters 2 and 3), but it can help us formulate entirely new questions that are philosophically interesting (chapter 4). However, not even all the data in the world can tell Richard what to do. Data can help Richard predict whether he will be praised or shunned for his decision, thus helping him build more attractive arguments, but it can never decide between John and George's parents. Thus, revisiting the normative debate that opened this introduction after considering

all the evidence reveals that Richard will likely remain puzzled. Nonetheless, a better understanding of rules will help him see more clearly which features should figure in his normative reasoning. Issues such as the recognitional community one has in mind might not be salient in decision-making absent experimental philosophy, but might – as we will see – nonetheless be normatively relevant. Finally, even if the practical guidance that an experimental account of rules can provide for judges is limited, such an investigation might contain clearer lessons for rule-makers: unlike judges, rule-makers often know who they address through legal rules. Thus, understanding how their addressees typically integrate text and purpose into rule violation judgments might help them make more efficacious rules.

2.
The empirical claims of the Hart-Fuller debate

As I have argued in the introduction, the fact that Hart and Fuller appeal to the intuitions of their audiences is evidence that their positions are predicated on certain empirical assumptions. Even if controversial by itself, this statement is silent on two fronts. First, which exactly are these positions? Second, what empirical evidence might bear on them? These two issues are potentially as controversial as the argument developed in the introduction about the importance of experiments to legal philosophy: after one accepts the idea that the appeals to intuition made by Hart and Fuller commit them to certain empirical claims regarding people's beliefs and attitudes, the interrelated questions of what precise claims are at stake and what kind of evidence is needed in order to test them remain open. This chapter is chiefly concerned with these questions, but before turning to them, it is important to acknowledge that the relevance of experimental evidence presupposes one specific interpretation of the task of general jurisprudence, one that is contested within Hartian scholarship. Moreover, it is important to clarify what is the role of exegetical claims in the overarching analysis pursued in this dissertation.

Hartians disagree about the right way to conceive of Hart's method. Some (e.g., Struchiner, 2005; Schauer, 2013a; Schauer, 2020; Nye, in prep) emphasize excerpts out of "The Concept of Law" where Hart expresses his doubts regarding the possibility and utility of finding a precise definition of law. According to them, Hart does not believe that the concept of law can be reduced to a set of individually necessary and jointly sufficient conditions. This, coupled with Hart's own self-professed goal of engaging in "descriptive sociology" (1994, p. v), leads to an

interpretation of Hart that focuses less on ambitious metaphysics and more on elucidating the many possible and actual ways in which human law manifests itself. According to this view, instead of identifying necessary truths about law, Hart is concerned with elucidating the *important*, even if not individually necessary, features of the phenomenon.

In contrast, another strand of contemporary analytic jurisprudence takes a different view of the Hartian enterprise. Philosophers such as Scott Shapiro and Joseph Raz are concerned with claims about what is necessarily true of law in all possible worlds (see Schauer, 2020; Nye, in prep). As Nye puts it

> […] many people (including Raz and Shapiro themselves) appear to consider Raz and Shapiro the torchbearers of a method pioneered by Hart. According to Shapiro, Hart "seemed to have assumed that the nature of law could be determined exclusively by conceptual analysis". Shapiro says that Hart's view, as he understands it, is right, and that it entails looking beyond human society. What we are looking to understand instead is "all possible instances of law" (in prep., p. 2).

Under this view, hypothetical thought experiments regarding alien worlds and societies of angels are just as relevant for general jurisprudence as actual data regarding law in this world. In contrast, knowledge about important but not necessary features of law falls within the realm of legal sociology, and not of legal philosophy (see the characterization of Raz in Schauer, 2020). As a consequence, empirical data about a finite set of rules is of only limited philosophical importance.

Even though this enterprise – which we may call metaphysical – might be interesting and fruitful in itself, I believe that the first reading is more faithful to Hart's own views expressed in his published works. Moreover, this is also the version of Hartian thought most responsive to empirical arguments. Hartians like Schauer and Nye – but maybe not those like Raz and Shapiro – are bound to revise their views about law in light of experimental results regarding the concept of rule. Since one of the crucial steps of the debate that interests me here involves appeals to intuitions that seem empirical in nature, this might be the best version of Hartian theory to explore in this specific context. Moreover this sort of anti-essentialist stance represents a currently influential take on the task of legal philosophy (see Priel, 2007). Thus, even if someone comes up with a conclusive exegetical argument to the effect that Hart was actually a metaphysician interested in modal

claims about all possible worlds, it would still be worthwhile to explore the concept of rule under Schauer's, Nye's and Priel's lenses.

What about Fuller? The scholarship on his work is much thinner[31], but also much less contentious. Recent works interested in rehabilitating Fuller's views argue that he was relegated to a lesser position by an undue focus among contemporary jurisprudents in the search for necessary and sufficient conditions of law. In contrast, "For Fuller, the relationship between law and morality should be analyzed not in terms of conceptual necessities but in terms of intrinsic presuppositions and empirical tendencies" (Burg, 2014, p. 739). Moreover, van der Burg argues that Fuller should be read as a pragmatist, "[…] interested not in conceptual necessities or rigid dichotomies, but in understanding how law works in specific contexts and for specific purposes" (p. 740). Thus, the non-essentialist, empirically committed view of the Hart-Fuller debate also seems to be the most promising reading on the Fullerian side of things.

This methodological issue might be the most controversial assumption underlying the exegetical claims advanced in this chapter. Nonetheless, it is not the only controversial statement I wish to make about the Hart-Fuller debate. Some philosophers would likely deny that Hart and Fuller held at least some of the positions I attribute them. In each case where I'm aware of possible pushbacks, I have tried to deal with them. But the interpretation I advance here is not meant to be especially polemic, as it finds support on most secondary literature about the Hart-Fuller debate (see the comprehensive reviews of the debate by Schauer, 2008 and Vega Gomez, 2014). Moreover, nothing special should turn on exegetical matters. The ideas and theses I will discuss should be interesting in themselves, even if they are misattributed to these authors. After all, rules are ubiquitous in practical reasoning: they figure prominently in legal argumentation, moral discourse, and even in less dramatic aspects of daily life, as exemplified by rules of etiquette and other norms of social behavior. Thus, a better understanding of what is and isn't true of rules should interest philosophers even when the claims under investigation lack any pedigree grounded in the history of legal philosophy.

[31] As noted by van der Burg (2014), "For most contemporary students of jurisprudence, Lon Fuller (1902-78) is hardly more than a footnote in the history of their discipline" (p. 736).

By its turn, answering the question of which kind of evidence is relevant to test their positions is hard because neither Hart nor Fuller had experiments in mind in 1958. Not only that, but legal philosophers have only recently started to conduct their own experiments, rendering even contemporary secondary literature of limited help. In any event, it is reasonable to assume that Hart and Fuller might have modulated their claims if forced to make empirical predictions. To account for the ambiguities in how to construe the appropriate evidence needed to support or reject a given view, I have attributed two theses to each author. The first pair of theses attributed to the authors make contrasting conceptual claims about possibility. As we will see, these claims are either too strong – requiring impossibly demanding confirmatory evidence – or too weak – requiring too little evidence and having low predictive power. The second pair of theses makes quantitative empirical claims requiring more reasonable evidence.

Which specific theses does this consideration of substantive positions and evidentiary requirements entail? I will argue that Hart might be read as defending two distinct theses in the relevant portion of his debate with Fuller. The first one is that at least some rules have a core of settled meaning that enables us to make rule violation judgments based only on the rule's text. The second thesis is that text is actually prevalent in ordinary interpretation: most people, when text and purpose clash, side with text. Fuller, on the other hand, counteracts each of those theses with a denial. First of all, Fuller claims that all rule violation judgments under all rules involve the moral purposes pursued by the rule. Second, he believes that, when text and purpose clash, most people (or at least just as many people) side with purpose.

I will then argue that the contrast between these theses is highly consequential for legal philosophy. One of the most famous divides in general jurisprudence is that between legal positivists and natural lawyers. Fuller's stronger thesis is incompatible with the central thesis of positivism. Finally, I will introduce Schauer's framework, complement it, and discuss how it can be leveraged to test the empirical presuppositions of the Hart-Fuller debate, preparing the terrain to review the existing evidence in chapter 3.

2.1
Four theses

Although Hart was clearly a textualist, his view about rules makes room for purposes to play an important role in some cases. As discussed in the introduction, Hart claims that the no-vehicles in the park rule's text is all we need to ascertain that regular cars are prohibited. However, in borderline cases involving skateboards, toy automobiles, airplanes, and many other things that may or may not count as vehicles, we must resort to other considerations, including purposes. Hence, Hart should not be read as denying that we sometimes resort to purposes in interpreting rules, but as affirming the prominent role of text in such enterprise:

> Perhaps the claim that it is wise [to drop the firm utilitarian distinction between what the law is and what it ought to be] cannot be theoretically refuted for it is, in effect, an *invitation* to revise our conception of what a legal rule is. We are invited to include in the "rule" the various aims and policies in the light of which its penumbral cases are decided on the ground that these aims have, because of their importance, as much right to be called law as the core of legal rules whose meaning is settled. But though an invitation cannot be refuted, it may be refused and I would proffer two reasons for refusing this invitation. First, everything we have learned about the judicial process can be expressed in other less mysterious ways. We can say laws are incurably incomplete and we must decide penumbral cases rationally by reference to social aims. [...] Second, to insist on the utilitarian distinction is to emphasize that the hard core of settled meaning is law in some centrally important sense and that even if there are borderlines, there must first be lines. [...] By contrast, to soften the distinction, to assert mysteriously that there is some fused identity between law as it is and as it ought to be, is to suggest that all legal questions are fundamentally like those of the penumbra. It is to assert that there is no central element of actual law to be seen in the core of central meaning which rules have, that there is nothing in the nature of a legal rule inconsistent with *all* questions being open to reconsideration in the light of social policy (Hart, 1958, pp. 614-615).

If there is something in the nature of a legal rule that is inconsistent with *all* questions being open to reconsideration in the light of social policy, then at least *some* rules must conclusively settle at least *some* questions definitively without resort to moral purposes. Hence, departure from a rule's text is sometimes sufficient for rule violation[32]. This claim might be restated as the Core of Settled Meaning Thesis (CSM):

[32] Brian Bix (2003) reads Hart differently. According to his reading, in *The Concept of Law*, Hart's claims were limited to legal interpretation and, as such, put special emphasis on the central cases the *legislator* had in mind. Hence, text would not be strictly sufficient by itself to entail that some cases

> Core of Settled Meaning Thesis (CSM): for at least some rules and some cases under those rules, textual violation is a sufficient condition for rule violation.

According to CSM, it is a feature of our concept of rule that, for the cases falling under the core of settled meaning of the rules that have one, violating the rule's text is enough to entail rule violation. Note that this is compatible with the fact that some prescriptions might be so vague that they may lack a core of settled meaning, however small[33]. Similarly, CSM doesn't deny that even clear rules might leave a lot of cases unsettled. In effect, the "no vehicles in the park" rule is fruitful precisely because it shows that clear rules also have a penumbra (where we don't know how to deal with "bicycles, roller skates, [and] toy automobiles", Hart, 1958, p. 607).

It should also be noted that, just as Hartian positivism, CSM is normatively inert (see Gardner, 2001; Schauer, 2011). It states that text is sometimes sufficient for rule violation, but says nothing about what should be done after one acknowledges that the rule was violated. CSM is compatible both with a formalist outlook that sees rule violation as dispositive (e.g., if the rule was violated, then the consequences it stipulates should always be brought about) and with a view that sees rule violation as only one among many factors that must be taken into account before deciding what should be done (e.g., if the rule was violated, but it prescribes an unjust result to this specific case, then we should side with justice). So, Hart's allegiance to CSM doesn't make him a formalist, and criticism against formalism does nothing against CSM[34].

Thinking about the evidence needed to support CSM, we can see that it is a rather weak thesis. Even presenting one single case where most people take text to be sufficient to entail rule violation would be enough. Hence, there is no *need* to engage in systematic empirical research to test CSM. Much less demanding methods could do the job. Regardless, systematic empirical research is nonetheless relevant. If we conduct several experiments and show that text is sometimes

violate the rule. Hart himself saw problems with Bix's interpretation, as disclosed by Bix's postscript to the relevant section of his book.

[33] But in that case, those prescriptions would escape the narrow definition of rule that I am employing here (see the discussion in the introduction debating Schauer, 1997).

[34] Similarly for other authors, such as Schauer (1991), who explicitly endorse CSM.

sufficient for rule violation, it is hard to deny that this supports CSM. Likewise, if, after many experiments, no one can come up with a case where text is indeed sufficient, our credence in CSM should decrease[35].

The effort needed to find empirical support for CSM is inversely proportional to the thesis' explanatory power. Saying that text is sometimes sufficient says nothing about how often this is the case. CSM is compatible with textual sufficiency being very rare (e.g., a fringe phenomenon affecting few rules and even fewer cases), and with a fully textualist account of rules (e.g., text is *always* sufficient to determine rule violation judgments). However, another possible reading of Hart, in line with the anti-essentialist thrust of authors such as Schauer and Nye, produces a more interesting, more controversial and more directly empirical claim. But before turning to it, it is important to investigate Fuller's denial of CSM.

Fuller's memorial truck example "meant to insist that it was *never* possible to determine whether a rule applied without understanding the purpose that the rule was supposed to serve" (Schauer, 2008, p. 1111). A couple of pages later, Fuller asks if it is "really ever possible to interpret a word[36] in a statute without knowing the aim of the statute?" (1958, p. 664), a question he wishes to answer with a negative. Hence, for him, rule violation judgments on *all* cases rely on purposes. I call this the Strong Purposivist Thesis (SP):

> Strong Purposivist Thesis (SP): rule violation is *always* a matter that needs to be settled by resort to moral purposes.

SP is a very ambitious thesis. It makes a universal claim about all rule violation judgments. If we can come up with one single case where moral purposes

[35] It is important to have in mind that the vocabulary of statistics is not perfectly well-suited to make claims as weak as CSM requires. Because of the possibility of measurement error, statistical analysis is not in the business of making assertions about what happens in one given case. Instead, the statistical approach involves quantifying our certainty that a hypothesis is true over a range of observations. The hypothesis might be quite weak – for instance, we might test whether people are any more likely to render rule violation judgments when text was violated relative to cases where text wasn't violated –, but it can't be quite as weak as philosophers would have it (where one single case could suffice). In what follows, in order to make sense of the statistical tests that were conducted, I will therefore treat CSM as being a bit stronger than it actually is. However, as we will see, nothing very important should hinge on these marginal differences in how to construe the evidence needed to confirm CSM.

[36] For the problems with Fuller's misguided characterization of Hart's position as turning into the meaning of words and not entire sentences, see Schauer (2008).

are irrelevant, SP fails[37]. Moreover, SP might be false even if CSM is also false. Suppose that the original intentions of the rule maker always play a role in the interpretation of rules in such a way that whenever text and original intentions clash, we privilege the original intentions (as argued by, among others, Alexander and Prakash, 2004; Alexander and Sherwin, 2008). In that case, CSM is false, because text is never sufficient to determine rule violation judgments, but SP is also false, since what makes it insufficient is the necessary resort to original intentions, and not to moral purposes.

Taking these features into account, the evidence required to support SP is especially strong: we need data showing not only that text is insufficient, but also that it is insufficient because of the necessary resort to moral purposes, and not for any other possible reason. Moreover, whereas one single counterexample might falsify SP, we would need an indefinitely great number of examples in order to accept it.

On the other hand, if SP were true, it would certainly be interesting and highly explanatory. So, whereas a lack of explanatory power was a shortcoming stemming out of CSM's modesty, it is not a concern with regards to SP. But SP's ambition is clearly problematic in a different way. If Fuller was truly concerned with making empirical claims about people's beliefs and attitudes regarding law, he should have been aware of the demanding burden of making a universal statement in that context. Perhaps, this reconstruction of SP focuses too much on one counter-example and some sentences that might have been meant hyperbolically[38]. So, just as a different reading of Hart might produce something stronger than CSM, a charitable reading of Fuller might require something weaker than SP.

[37] Hart's original "regular car" example is unfit here because, as Fuller rightfully points out, this is a case where both text and purpose recommend the same result: "If the rule excluding vehicles from parks seems easy to apply in some cases, I submit this is because we can see clearly enough what the rule 'is aiming at in general' so that we know there is no need to worry about the difference between Fords and Cadillacs. If in some cases we seem to be able to apply the rule without asking what its purpose is, this is not because we can treat a directive arrangement as if it had no purpose. It is rather because, for example, whether the rule be intended to preserve quiet in the park, or to save carefree strollers from injury, we know, 'without thinking', that a noisy automobile must be excluded" (1968, p. 663).

[38] See Rundle, 2012, p. 6, stating that "numerous instances of hyperbole [infected] some of his claims".

What would those variations look like? Starting once again with Hart, we can shift the emphasis to the idea that there is something "centrally important" to law about textualism. If textual sufficiency is centrally important, than Hart should be read as committed to something more ambitious than CSM. After all, CSM is indifferent between textual sufficiency happening often or extremely rarely, but in order to be centrally important to law, textualism should be a pervasive feature of the interpretation of rules. I call this stronger, more clearly empirical Hartian claim the Textual Prevalence thesis (TP)[39]:

> Textual Prevalence thesis (TP): text matters more than purpose when assessing rule violation.

There are at least two plausible readings of TP that entail different empirical predictions. The first way understands TP to be a claim about the extension of the core of settled meaning. Under this reading, TP states that there are more cases falling under the core of settled meaning than under the penumbra. This is certainly an idea Hart subscribed to. He made constant references to the fact that the penumbra was something that happened "at the borderline", while the core of settled meaning covered "the great mass of ordinary cases" (Hart, 1994, p. 128)[40]. While interesting, explanatory, and empirically testable, it might be hard to produce empirical tests for this version of TP. To do so, one would need to identify a representative sample of rules and characterize every instance of rule violation or compliance as belonging to the penumbra or to the core of settled meaning.

The second reading of TP makes a claim about what happens when text and purpose clash under the core of settled meaning: text wins. Most of the times when the rule's text is violated, but its purpose isn't, most people say that the rule was violated. In contrast, whenever a rule's text is not violated, but its purpose is, TP claims that most people in most cases would say that the rule was not violated. Here, I will focus on this second reading. In order to test it, it is necessary to identify some situations where text and purpose disagree and see which side is privileged more often.

[39] On that point, see Schauer, 2008, p. 1130.
[40] See more generally Hart, 1994, chapter VII. See also Struchiner (2002).

TP might provide a way to remedy Fuller's hyperbole: perhaps, it is charitable to construe him as simply denying TP. I call this alternate purposivist view the Weak Purposivist thesis:

> Weak Purposivist thesis (WP): purposes matter at least just as much as text when assessing rule violation.

Even though it is harder to find direct support for this thesis on Fuller's writings[41], it is very much in the spirit of his scholarship. Understood as the negation of TP, WP makes claims about *clear* cases. If even on clear cases under the core of settled meaning, purposes are just as important as text, then the Fullerian emphasis on purposes is vindicated.

One last point regarding exegesis must be discussed before moving on. In the introduction to *Essays in Jurisprudence and Philosophy*, Hart stated the following:

> In fact, as I came later to see and to say in Essay 4 [*American Jurisprudence Through English Eyes: the Nightmare and the Noble Dream*], the question whether a *rule* applies or does not apply to some particular situation of fact is not the same as the question whether according to the settled conventions of language this is determined or left open by the words of that rule. For a legal system often has other resources besides the words used in the formulations of its rules which serve to determine their content or meaning in particular cases. Thus, as I say in Essay 3 [*Problems of the Philosophy of Law*], the obvious or agreed purpose of a rule may be used to render determinate a rule whose application may be left open by the conventions of language, and may serve to show that words in the context of a legal rule may have a meaning different from that which they have in other contexts. My failure to make this clear amounts, as Fuller argued in his reply to Essay 2 [the references are to Fuller (1958) and Hart (1958)], to a defective theory of statutory interpretation, which I do something to correct in my later discussion, in Essay 3, of what it is that makes clear cases clear. Certainly my arguments need to be both amplified and corrected on this point. But their correction does nothing to support the claim that the law is a gapless system never incomplete and containing within it an answer to every question of law even in the hardest of hard cases […]. (1983, pp. 7-8).

Here is the relevant passage of Essay 3 (*Problems of the Philosophy of Law*):

> Rules cannot claim their own instances, and fact situations do not await the judge neatly labelled with the rule applicable to them. Rules cannot

[41] Sometimes, Schauer seems to attribute this sort of claim to Fuller: "Fuller's point [with his counterexample to the 'no vehicles in the park' rule] was that something, the statue of a vehicle, might be within the letter of a rule but yet not within its purpose, in which case the letter should yield to the purpose" (1991, p. 74). Overall, though, it is clear that Schauer reads Fuller as committed to SP.

> provide for their own application, and even in the clearest case a human being must apply them[42]. The clear cases are those in which there is general agreement that they fall within the scope of a rule, and it is tempting to ascribe such agreements simply to the fact that there are necessarily such agreements in the use of the shared conventions of language. But this would be an oversimplification because it does not allow for the special conventions of the legal use of words, which may diverge from their common use, or for the way in which the meaning of words *may be clearly controlled by reference to the purpose* of a statutory enactment which itself may be either explicitly stated or generally agreed (p. 106, emphasis added).

Is this a retraction of CSM? Perhaps yes, but it would be a very puzzling retraction. First, the idea that words might have special legal meanings poses no problem for the thesis. CSM might presuppose that a text's meaning is a matter of convention, but it doesn't specify *whose* convention[43]. A convention among legal professionals to use word W, with ordinary meaning O, with the meaning O_L in a legal context is sufficient to ensure the viability of CSM. Suppose that lawmakers privy to the convention enact a rule stating: "No Ws allowed". A judge could surely recognize that a case falling clearly within the scope of O_L (but outside the scope of O, for instance) is clearly in violation of the rule. No resort to purposes of any kind is necessary to do so. Moreover, CSM specifies only that *some* rules might have a core of settled meaning. So, even those who reject this rejoinder to the objection would still be forced to acknowledge that CSM might still hold with regards to those rules that employ only ordinary words with their ordinary meanings. Second, the concession that the meaning of words *may* be controlled by reference to purpose is compatible with CSM. To deny CSM, one needs to say something stronger, something to the effect that words *have* to be controlled by purposes (SP), for instance.

One way to make sense of the passages above while retaining CSM is to understand Hart as defending himself against charges of formalism[44]. However,

[42] So far, nothing new. See how similar the language is to *Positivism and the Separation of Law and Morals*: "The toy automobile cannot speak up and say 'I am a vehicle for the purpose of this legal rule', nor can the roller skates chorus, 'We are not a vehicle'. Fact situations do not await us neatly labelled, creased, and folder; nor is their legal classification written on them to be simply read off by the judge" (1958, p. 607).

[43] About the way in which a word's meaning is always contextually dependent and how Hart's doctrine of the CSM is perfectly compatible with this, see Schauer (2008). Elsewhere (Almeida, 2017), I have attempted a more thorough characterization of how the conventions of legal language relate to ordinary language and the normative implications of this relationship for legislators.

[44] See the last part of the first quote: "Certainly my arguments need to be both amplified and corrected on this point. But their correction does nothing to support the claim that the law is a gapless

CSM would remain interesting no matter what Hart thought of it. This, in conjunction with his earlier stance, is enough to warrant CSM's inclusion in this investigation.

2.2
Textualism, purposivism and legal positivism

In the last section, I have argued that Fuller was a proponent of SP and WP, while Hart defended CSM and TP. But why are those contrasts important? Do they imply anything deeper, or are they just theoretical disagreements about legal interpretation? I believe that at least one of those contrasts – the one between CSM and SP – is important because of the broader implications of the theses under contention. This happens because SP is something that Hart cannot concede to as a legal positivist[45]. Famously, there is disagreement about what 'positivism' entails, but all self-identified positivists accept the following proposition:

> Separability Thesis (ST) For a normative system to be called legal, or for a rule to be called legal, no moral test or criteria must be met (adapted from Struchiner, 2005, p. 33)[46].

It might be useful to return to the 'No vehicles in the park' example to showcase the tension between CSM and SP with regards to ST. Suppose that the public park where vehicles are prohibited is the object of a local statute. This statute directs the holder of the office of director to regulate what is to be allowed inside the park. Using his powers, the director affixes a sign on the park's entrance that reads 'no vehicles allowed'. ST seems to require the possibility that in some legal systems, people are able to infer that vehicles are prohibited in the public park without resort to moral reasoning. Put differently: knowledge of social facts, under

system never incomplete and containing within it an answer to every question of law even in the hardest of hard cases […]" (Hart, 1983, p. 8).

[45] But see Bix: 'Hart would accept Fuller's position as the starting-point for judicial interpretation, but he would disagree that such resources are sufficient to avoid the need for judicial discretion' (2003, p. 29).

[46] Struchiner, by his turn, took inspiration from Gardner (2001). Gardner's version of the thesis states that "In any legal system, whether a given norm is legally valid, and hence whether it forms part of the law of that system, depends on its sources, not its merits" (p. 199). Gardner goes on to elucidate that "'Source is to be read broadly such that any intelligible argument for the validity of a norm counts as source-based if it is not merits-based" (p. 200). Struchiner's emphasis on particular legal rules and direct reference to moral criteria makes the argument that I wish to make much clearer than Gardner's more popular (but perhaps also more truncated, with its stipulative use of "source") formulation.

some possible legal systems, needs to be sufficient to arrive at the legal norm for ST to obtain.

For positivists, some of these social facts could be formulated as the following statements:[47]

> (1) that there is a statute that was enacted according to the appropriate lawmaking procedures;
>
> (2) that this statute establishes the office of park director and makes the holder of the office responsible for saying what can and cannot be allowed inside the park;
>
> (3) that the current holder of the office of director, appointed pursuant to law, affixed a sign to the park's entrance which read 'no vehicles allowed';
>
> (4) that the sign means that some things - for instance, a regular car - are clearly prohibited from the park.

SP takes no issue with the factual status of statements 1-3, but it challenges the idea that we can get to statement 4 purely by virtue of statement 3, plus factual linguistic conventions. If we always need to resort to morality in order to determine what the law requires even in linguistically clear cases, then morality plays a necessary role in determining the content of legally valid rules. Put differently, the core commitment of positivism requires that at least some rules might be identifiable without reference to their moral standing. Insofar as comprehending the requirements of a rule can be a moral-free enterprise, ST is safe, but if rule comprehension *requires* the comprehension of moral purposes (SP), positivism fails.

CSM is one way to make the jump from (3) to (4) merits-free[48], but it is not the only one. So, denial of this thesis is not sufficient to make someone a natural lawyer. When Alexander and Prakash say that '[intention free] textualism is a conceptual impossibility' (2008, p. 969), they are also denying CSM. But instead of postulating the need to resort to moral purposes (SP), they insist that, in order to understand a rule, we need to capture the intentions of the rule-maker. Since the

[47] This list could have been much longer: Hart (1994) uses the strategy of spelling out the entire list of rules one needs to know in order to say that a rule is part of a legal system to show that the system rests on the ultimate rule of recognition. At this point, however, we don't need to go that far.

[48] As acknowledged by Fuller himself: "Can it be possible that the positivistic philosophy demands that we abandon a view of interpretation which sees as its central concern, not words, but purpose and structure? If so, then the stakes in this battle of schools are indeed high" (1958, p. 667).

actual intentions of legislators count as historical data, the authors can maintain their position while still adhering to the positivist thesis ST[49]. What makes SP a natural law thesis is hence not the denial of CSM, but the idea that there is something necessarily moral in interpretation.

To sum up, I discussed two pairs of contrasting theses. The first pair contrasts CSM, the idea that text is sometimes sufficient for rule violation, with SP, which states that all rule violation judgments require the consideration of the rule's moral purposes. CSM is weak and requires very little data to be vindicated. If we can find only one unequivocal case where most people are willing to say that text is sufficient, even though the rule's purpose indicate otherwise, that is enough to justify credence in CSM[50]. In contrast, SP is extremely strong and requires a wealth of evidence to be vindicated, and very little data to be falsified. A single counterexample would suffice. Moreover, the contrast between CSM and SP is highly impactful. SP is incompatible with one popular way to construe legal positivism (ST). Thus, knowing whether SP is true is important.

The second pair contrasts theses of comparable empirical ambition. On the one hand, TP argues that, when text and purpose clash, text wins most of the times. On the other, WP denies that thesis, stating that purpose is at least as prevalent as text in determining people's rule violation judgments.

2.3
Devising a test

How, exactly, can we test these four theses? Which cases should we consider? Schauer's (1991) inquiry into the nature of rule provides a framework to design experiments aimed at testing the roles of text and purpose in the concept of

[49] It is also important to emphasize that Fuller is not a strong intentionalist. In his reply to critics of *The Morality of Law* he explicitly criticizes the idea that law might be interpreted solely by reference to the intentions of a lawgiver. See Fuller, (1969), 228-229. I will come back to this discussion on chapter 5.

[50] Which people counts is once more debatable: if we restrict the recognitional community to judges, then it is sufficient to find a case to which most judges are willing to say that text is sufficient.

rule. To illustrate this framework, let us consider his favorite example: that of a rule prohibiting dogs in a restaurant (p. 25).

Imagine that the owner of a restaurant has several bad experiences with costumers who bring their dogs with them. The dogs often misbehave. They bark and snarl at other costumers, making their experience unpleasant. Aiming to avoid nuisances to his clients, the owner creates a rule with the following text: "No dogs allowed in the restaurant".

Now, imagine that an eccentric woman takes her pet tiger to the restaurant. Is she in violation of the rule? The rule's text unequivocally allows the tiger, because tigers are not dogs. In contrast, the moral purpose pursued by the rule clearly bars the tiger. After all, tigers can cause much more trouble than dogs. What goes for tigers also goes for bears, snakes, and drunk people. Each of these cases falls comfortably outside the textual scope of the rule, but each is also prohibited by the rule's purpose. Schauer calls this class of cases *underinclusion*: the rule's text includes less than it should (when evaluated by the moral purpose it is supposed to advance).

Contrast this set of cases with this situation: "[a] child enters the restaurant carrying a purse containing what seems to be a teddy bear. Actually, it's her dog, which doesn't bark and barely moves, being easily mistaken for a toy" (Struchiner, Hannikainen & Almeida, 2020). Here, the text unequivocally bars entrance: we are dealing with a dog. But this dog is so quiet and cute that no one would ever be bothered by it. This situation characterizes what Schauer calls *overinclusion* cases: the rule's text includes more than it should.

Even though Schauer stops here, there are two other classes of cases defined by the relations between text and purpose: *core* cases, where both text and purpose agree that something should be included within the rule's scope (for instance, loud, misbehaving dogs), and *off-topic* cases, where neither text nor purpose apply (a woman walks into the restaurant with a pet goldfish on a plastic bag full of water).

If we assume that the lay concept of rule is defined by its use, *core*, *overinclusion*, *underinclusion*, and *off-topic* cases can be leveraged to test CSM, SP, TP and WP. While all theses agree that people will say core cases are (and off-

topic cases aren't) rule violations, they make competing predictions about over- and underinclusion cases. Specifically, CSM states that at least some cases of overinclusion will be seen as rule violations by a majority of people, while SP predicts the opposite (no cases of overinclusion will be seen as rule violations)[51]. Moreover, SP also predicts that most people will take underinclusion cases to be rule violations. TP and WP, on the other hand, make comparative predictions regarding over- and underinclusion cases. TP states that overinclusion cases are seen as rule violations more often than underinclusion cases, while WP claims that underinclusion cases will be thought to violate the rule at least as often as overinclusion cases.

[51] See the previously raised caveats above about the use of statistical methods to test strong conceptual claims.

3.
The evidence

In the last chapter, I discussed the theses advanced by Hart and Fuller in their debate about the concept of rule. For each thesis, I considered what kind of evidence would be needed to support it. Finally, I argued that Schauer's framework provided a tool to gather that evidence[52].

That is exactly what we tried to do in "An Experimental Guide to Vehicles in the Park" (Struchiner, Hannikainen & Almeida, 2020). This chapter begins with a review of the four experiments reported in that paper and how they relate to the theses. The studies involved lawyers and laypeople, as well as legal and non-legal rules and provide important evidence bearing on the Hart-Fuller debate. Overall, they show that people are textualists when prompted to consider text and purpose before assessing whether a rule was broken, but that purposes are very important in informing people's judgments (lawyers and laypeople alike) when the only question asked concerns rule violation. One way to make sense of these results is to conceive the concept of rule as composed of two distinct components: one factual (text), and the other moral (purpose).

However, that is not the only interpretation possible. To show that, I will review work done by Turri & Blouw (2015) on rule violations. They argue that rule violation judgments entail a judgment of blameworthiness. If participants in our experiments are using their answers to signal not only their concept of rule, but also

[52] Earlier works about rule violation judgments such as Garcia, Chen & Gordon (2014), Turri & Blouw (2015) and Turri (2020) that did not use Schauer's categories as a starting point are incomplete with regards to his categories. Either they don't pitch text and purpose against one another (as is the case with Turri & Blouw, 2015; Turri, 2020), or focus solely on overinclusion cases (Garcia, Chen & Gordon, 2014)

their judgments of blame, then rule violation judgments might not be the best way to test the Hart-Fuller debate. After all, purposive responses could be nothing more than an expression of blame attribution. For this reason, we conducted another study, this time looking at rule identity judgments (that is, about whether rules changed or not after alterations in text and purpose. See Almeida, Struchiner & Hannikainen, under review). The results revealed a difference between the way lawyers and laypeople employ the concept of rule: lawyers' rule identity judgments were entirely textual, while lay people's judgments were also informed by purpose.

What are the consequences of this difference between lawyers and laypeople to the Hart-Fuller debate? The fact that lawyers and laypeople have different concepts of rule might go a long way in explaining some of the differences between the two philosophers. Hart focused on the role of officials (e.g., Adler, 2006), while Fuller was emphatic in including citizens in his account (Rundle, 2012). Perhaps, both were fundamentally right about what goes into the rule violation and rule identity judgments of lawyers and laypeople, but their disagreement about the appropriate scope and emphasis of legal theory (in particular, about the recognitional community) might have led to wider disagreement about philosophical claims in law.

In what follows, for ease of reference, I will number studies sequentially as I present them. Since I am going to cover all four experiments contained in Struchiner, Hannikainen & Almeida (2020) first, these four studies will receive the same numbering as they did in the original paper. This, however, is not the case for the studies I will discuss from Turri & Blouw (2015)[53]. That paper contains seven different experiments, and only two of them (Experiments 2 and 4, in the original numbering, Studies 5 and 6 here) will be discussed. Finally, Almeida, Struchiner & Hannikainen (under review) reports only one experiment, which I am going to refer to as Study 7. Later on, chapters three and four will also follow the convention, presenting further studies from papers currently in preparation.

[53] To be clear: I have nothing to do with this paper. It just happens to report results that are very important for my overall argument, so much so that it is useful to introduce some shorthand to refer to them.

3.1
An experimental guide to vehicles in the park

3.1.1
Study 1

In true Schauerian fashion, Study 1 focused on the "no dogs in the restaurant" rule. At the beginning of each survey, participants read the following description of the story leading to the adoption of the rule:

> One day, a black dog called Angus ran, jumped around, barked and ate off the floor in a restaurant. Such case was thought to be the paradigm of something to be avoided in the future: behaviors that cause nuisances to costumers. Thus, the restaurant's owners created a rule: "no dogs in the restaurant"[54].

The story was meant to convey the rule's text ("no dogs in the restaurant"), as well as its purpose ("to avoid nuisances to costumers"). Subsequently, participants read four vignettes, randomly selected from a total of eight. In each vignette, a person entered the restaurant bringing something with her that varied across four dimensions: being a dog (which tracks directly the rule's text), looking like a dog, behaving like a dog, and annoying other customers (corresponding to the scope of the rule's purpose).

In some versions, the thing was a misbehaved dog that "runs, jumps, barks and eats from the floor", while other versions described quiet, unmoving dogs ("a child enters the restaurant carrying a purse containing what seems to be a teddy bear. Actually, it's her dog, who doesn't bark and barely moves […]"). In yet another set of versions, the thing was not a dog. Sometimes, it annoyed other customers ("a kid enters the restaurant with a cutting-edge toy: an extremely realistic robot dog, identical to a real dog and who acts like a real dog: it barks, jumps, drools, and walks on four paws"), while other times, it did not ("a kid enters the restaurant carrying a goldfish inside a plastic bag filled with water"). We also included some colorful examples that seemed to us as borderline examples for some of the dimensions (e.g., "A man enters the restaurant bringing a taxidermied dog, which he takes everywhere he goes […]").

[54] All Stimuli, Data, and analysis scripts for "An Experimental Guide to Vehicles in the Park" are available at: https://osf.io/7ft6e/

After each vignette, participants answered five different questions. The first one was our dependent variable: whether the person broke the rule while entering the restaurant. Possible responses were simply "yes" or "no". Afterwards, we asked each participant to assess whether the animal/object: (i) was a dog, (ii) looked like a dog, (iii) behaved like a dog, and (iv) bothered other customers. Here, possible responses ranged from 1 ("Clearly not") to 7 ("Clearly yes").

We were looking for correlations between the four assessments and the rule violation question. For each of the assessments, if it represents an element of the concept of rule, higher scores should be associated with rule violation judgments, and lower scores should accompany rule compliance judgments.

The theses surveyed in the preceding chapter make predictions with respect to two of the assessments. CSM predicts that (i) – whether the animal/object was a dog – correlates significantly with rule violation judgments, while TP claims that that correlation will be higher than the one between (iv) – whether the animal/object bothered other customers – and the question probing whether the rule was broken. On the other hand, SP entails a significant correlation between (iv) and answers about rule violation, while WP predicts that (iv) will correlate at least as strongly with rule violation judgments as (i).

Our results, gathered from 272 volunteers, including 45 persons who reported at least some legal training[55], showed that both assessments (i) and (iv) correlated with rule violation judgments, with (i) having a larger effect. Assessments (ii) and (iii), on the other hand, did not correlate with our rule violation question.

The results conform to the predictions of CSM. If text is sometimes sufficient, the judgment of whether something counts as a dog should at least sometimes be decisive in causing rule violation judgments. That entails a

[55] We asked participants whether they (a) were currently law students, (b) graduated from law school, or (c) neither. In what follows, I will treat everyone who gave answers (a) or (b) as "lawyers". Naturally, this glosses over many important and potentially influential differences between students and practitioners. However, our exploratory analyses failed to show any meaning full differences between lawyers and law students in the dependent variables of our studies. In any event, future work should drill down on the populational differences we discovered with these distinctions in mind.

correlation between (i) and rule violation judgments. Clearly, showing such a correlation is not sufficient to establish CSM. After all, this analysis is silent about what happens when purpose disagrees with text. However, had we found no correlation, we should diminish our credence in CSM.

The picture is slightly more complicated with SP. As we have seen, SP makes a universal claim about the role of purposes in all cases. In Study 1, we did not control the vignettes in such a way as to make any direct assessment of SP. However, if SP is true, than we should see a correlation between (iv) and rule violation judgments: assuming that purposes matter in every instance, cases that are seen as more clearly violating the purpose should be judged to violate the rule. Hence, even though the presence of a significant correlation does not help a lot in establishing SP, its absence could potentially speak against it.

Moreover, the correlation between textual assessments and rule violation judgments were higher than that between purposive assessments and the dependent variable. This result entails a straightforward implication for the less extreme pair of theses in the debate, supporting TP over WP.

In any event, the discussion of the implications of the results for each thesis makes clear the need for further experiments employing a more rigorous method, one that is able to provide more clear tests of the theses. That was the goal of Study 2.

3.1.2 Study 2

Study 2 leveraged Schauer's categories in a more direct way. We included four conditions, each corresponding to one of the four categories defined by Schauer's analysis covered in the last section of the preceding chapter: core, overinclusion, underinclusion, and off-topic cases. For each condition, we wrote four vignettes, each describing a scenario under one of four rules: "no vehicles in the park", "no shoes in the house", "no sleeping at the station"[56], and "no cellphones in class". Once more, in every scenario, the rule's text (e.g., "no cars are allowed

[56] This set of cases were inspired by Fuller's examples (1958, p. 664), but adapted to emphasize over- and underinclusiveness.

inside the park") and its purpose ("to avoid accidents that can imperil park goers physical integrity") were made explicit.

We then designed the survey in such a way as to ensure that each participant answered questions under each of the four conditions and each of the four rules. For instance, a participant could first receive questions about an overinclusion case under the "no vehicles in the park" rule, then about an underinclusion case under the "no cellphones in class" rule, followed by questions regarding a core case under the "no sleeping at the station" rule, and, finally, questions about an offtopic case under the "no shoes in the house" rule.

Once more, our dependent variable was a simple yes-or-no question about whether the person broke the rule. We also included assessments of violations of text (e.g., "Did the person entered the park with a car?") and purpose (e.g., "Did the person increase the risk of an accident?") in order to validate that the cases we saw as pertaining to each of Schauer's categories were also seen this way by our participants (in other words, these additional questions served as a manipulation check).

Up front, it is possible to see that this provides a better test for the four theses of the Hart-Fuller debate. CSM, as a claim about *sufficiency*, makes unique predictions in cases where *only* text – but not purpose – recommends rule violation. In other words, overinclusion cases provide the best test for CSM. If most people take at least some overinclusion cases to be instances of rule violation, this supports CSM. On the other hand, if a majority consistently answers "no" to the rule violation question in cases of overinclusion, this runs counter to CSM. As formulated, CSM makes no claim about underinclusion cases. Put another way, CSM is about sufficiency and not about necessity. In contrast, SP makes predictions about both over- and underinclusion cases. First, it predicts that people will overwhelmingly judge overinclusion not to be in violation of the rule. Second, if reference to moral purposes is centrally important for rule violation judgments, then underinclusion cases should also be seen as rule violations more often than not. Turning to the second pair of theses, we can see that TP makes a comparative claim between over- and underinclusion cases. More precisely, it predicts that people are going to judge overinclusion cases to be in violation of the rule *more often* than

underinclusion cases. On its turn, WP makes the claim that underinclusion cases are going to be perceived as rule violations at least as often as overinclusion cases.

Figure 1 summarizes the results of our 200 volunteers (98 of whom reported at least some legal training)[57]:

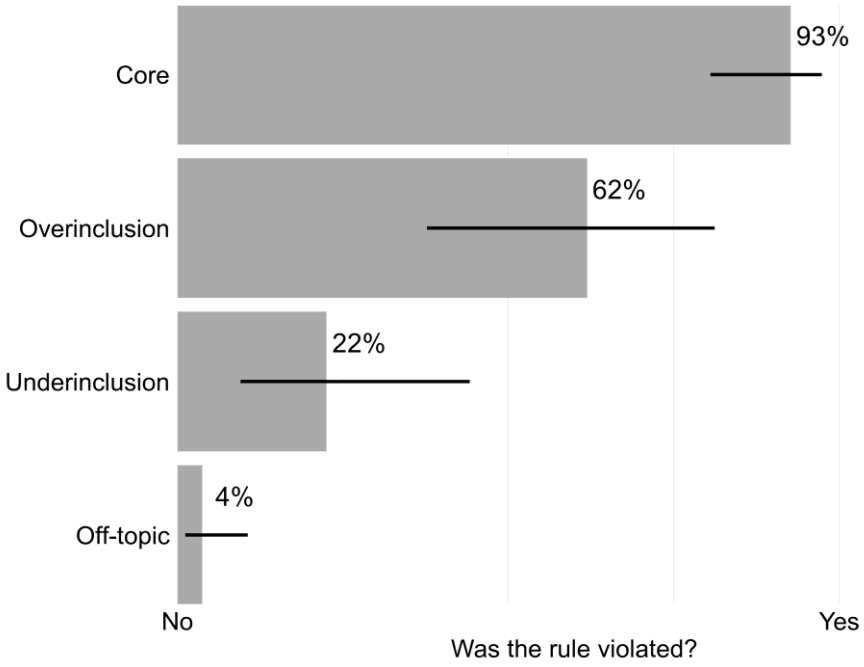

Figure 1 - A majority of participants judged overinclusion cases to violate the rule. In contrast, only 22% of participants judged underinclusion cases to be in violation of the rule[58]. Struchiner, Hannikainen & Almeida (2020, p. 319).

A majority of participants deemed overinclusion cases to violate the rule, while only a slim minority made the same judgment in underinclusion cases. Reassuringly, an overwhelming majority of our participants considered core cases to be - and off-topic cases not to be - in violation of the rule. This overall pattern obtained in three of the four scenarios[59], and the results were very similar when we

[57] In Struchiner, Hannikainen & Almeida (2020), this same figure is identified as Figure 2.

[58] Bars represent the predicted probability of a rule violation by condition. This means that the proportions shown are not simply those that obtained under each condition, but also takes into account the effects of scenario (which rule was discussed on that occasion?) and demographic variables. In any event, the implications would be the same if we used simple descriptive statistics instead of the more complex route we took in Struchiner, Hannikainen & Almeida (2020).

[59] Ironically, our putative overinclusion condition under the "no vehicles in the park" rule, which reads "A kid enters the park and plays with his toy car" was overwhelmingly seen to comply with the rule. Most participants did not think that a toy car counted as a car, as was confirmed by an analysis of the textual violation question. As such, it is hard to draw any conclusions based on that specific scenario. In any event, the fact that the pattern of results looks as strong as it does while including an off-topic case under the overinclusion moniker means that, if anything, we are

restricted our analysis to subjects with legal training. They support both CSM (a majority of participants thought that overinclusion cases violated the rule) and TP (overinclusion cases were judged to violate the rule more often than underinclusion cases), while casting doubt on SP and WP.

3.1.3
Study 3

Studies 1 and 2 dealt with clear cases. Their vignettes described cases that were either clearly infractions of the rule's text and/or of the rule's purpose. But one could argue that those cases are not particularly prevalent in the real world, or at least not among those cases that gather the attention of courts. Hence, we should also test cases that are less clear both with regards to being violations of text and of purpose.

When introducing TP, I argued that one of Hart's testable assumptions was that the core of settled meaning was larger than the penumbra. In other words, Hart would likely disagree with those objections to Studies 1 and 2. Nonetheless, inasmuch as TP is independent of this other assumption, testing whether text is still decisive even in more contested cases is still fair and informative.

To widen our gaze, we, in conjunction with undergraduate students working with the NERDS research group, came up with 24 vignettes that deliberately included mostly borderline cases such as the following:

> One day, in a high profile case, a 21 year old young woman got into a deadly traffic accident. The accident happened because the young woman was using her smartphone in one hand to text her friends while driving. Congress, recognizing the graveness of the situation and with the goal of avoiding this type of accident, passed a law with the following textual formulation: "it is forbidden to send text messages while driving".
>
> Felipe uses the voice-to-text functionality of his smartphone to text his friends. While doing so, Felipe suffers a serious accident with another vehicle, severely injuring the occupants of both cars.

downplaying the effects of text. Further evidence that the results might underestimate the effects of textual violation can be found in Garcia, Chen & Gordon (2014). The second study reported in that paper shows that 87% of participants judged the rule to have been violated in an overinclusion case.

The vignette reproduced above, while still explicitly stipulating the rule's text ("it is forbidden to send text messages while driving") and purpose ("avoiding this type of accident"), describes a case in the penumbra. If, on the one hand, Felipe *is* sending text messages while driving, on the other, he is doing so through his voice in a way that is much more similar to having a casual conversation with a passenger in your car than to the prototypical case of texting, where your hands and eyes are involved in a more distracting way.

Other vignettes were designed to be ambiguous with respect to purpose. For instance:

> Due to an increase in the persecution of religious minorities, the Department of Transportation budgeted for the installation of ecumenical chapels in all airports. Each chapel is clearly identified with a sign that reads "Members of all religious traditions are welcome to use this chapel to pray. Please, do not use this space for other purposes".
>
> James, who is from a predominantly catholic country, was harassed in the past for his agnosticism. He recently started reading "The Numerical Fabric of the Universe", the biography of an atheist mathematician who won the Nobel prize. James is constantly interrupted and hindered in his reading by his religious colleagues. Thus, in between flights, James goes to the ecumenical chapel in order to peacefully read his book.

Here, it is unclear whether the harassment James is avoiding counts as the persecution of a religious minorities. Atheism is not a religion, but atheists could conceivably be persecuted for their beliefs about religion. Hence, it is hard to classify this vignette into one of the Schauerian categories.

Study 3 was also aimed at another worry arising from Studies 1 and 2. In both of them, rule violation judgments were rendered in conjunction with assessments of textual and purposive compliance. It was plausible that this feature of the experimental design primed participants into taking these two specific factors into account in a way that they normally wouldn't. After all, subtle manipulations regarding joint and separate evaluation (e.g., Struchiner, Almeida & Hannikainen, 2020; Donelson & Hannikainen, 2020), and even the order with which cases are presented to participants (Schwitzgebel & Cushman, 2012) have been shown to influence participants' responses.

To address that, in Study 3 each participant was randomly assigned to answer only one question (either about the violation of the rule, its text, or its

purpose) about a random sample of four out of the 24 possible cases. In the rule violation condition, participants answered only whether or not each case violated the rule (e.g., "James broke the rule stated in the sign")[60]. In the text condition, the question asked dealt with textual violation (e.g., "James is a member of a religious tradition and is using the chapel to pray"), while in the purpose condition, the question dealt with moral blameworthiness[61] (e.g., "James should be morally chastised for reading the atheist mathematician's biography inside the chapel").

This experimental design allows us to see how people spontaneously integrate information about text and purpose into rule violation judgments. By averaging, for each case, independent judgments made by different people for each dimension, we can measure the influence of text and purpose over rule violation judgments. At the same time, in this design, participants aren't in any way directed to consider these two features in particular, instead of exclusively one of them, or even something else entirely, when rendering rule violation judgments.

What predictions are entailed by the theses under consideration? Here, CSM is arguably silent. As a thesis about what happens under the core of settled meaning, it simply is not relevant. On the other hand, SP – which applies more broadly to every rule violation judgment – predicts a significant case-level correlation between purpose and rule violation assessments. TP and WP, by their turn, make claims about the magnitude of the correlations. TP predicts that the higher correlation will be the one involving text, while WP states that the correlation between rule and purpose judgements will be at least as high as the first one.

175 persons participated (132 reported some legal training). Strikingly, in Study 3, the correlation between purpose and rule violation was slightly *higher* than that between text and rule violation. In other words, unlike in Studies 1 and 2, where TP prevailed over WP, in Study 3 the reverse is true, with WP coming out better.

[60] This time, rule violation judgments were recorded on a 7 point Likert scale similar to the one used in the assessments of text and purpose on Studies 1 and 2.

[61] This is another potential difference between Studies 1 and 2 and Study 3. In the first two cases, the question about purpose was formulated more narrowly around the specific goals pursued by rule-makers. Here, the broader question might pick out a different set of concerns. I return to this point in chapter 5 when discussing the precise meaning of purpose in legal judgements.

Again, the results are still the same when we restrict our analysis to those who received legal training.

Which of the differences introduced by Study 3 caused this reversal? This question is highly consequential.

Suppose first that the reversal was caused by the shift to penumbral cases. In that case, it would be hard to take any lessons from Study 3 to the dispute between CSM and SP. After all, CSM is irrelevant in borderline cases; and SP is polemical precisely in positing that resort to purposes is necessary even when text is clear. On the other hand, if WP obtains in borderline cases, but TP is true for clear cases, we would have a precise picture of the conditions under which each thesis fares better (one very similar to that held by Hart).

Now suppose that it was the choice to present each question in isolation that caused the difference. This finding could potentially bear on CSM and SP. If this is the case, then it was the introduction of questions highlighting both text and purpose that caused participants to lean Hartian in Studies 1 and 2. The lesson for the clash between TP and WP is also quite different under this assumption. If what makes people privilege text or purpose is not the fact that cases lie in the penumbra, but the way they engage with the question, then the conditions under which each thesis is true are completely different.

Study 4 represents our attempt at disentangling these two distinct explanations for the differences between Studies 1 and 2, on the one hand, and Study 3, on the other.

3.1.4
Study 4

To determine whether the questions we ask prime participants in some way, we designed an experiment where each condition introduced a different set of questions. Consider the following vignette:

> A 21 year-old woman suffered a traffic accident that took her life. The young woman was driving under the influence. In light of this tragedy, in order to avoid future accidents, Congress passed a zero tolerance policy, establishing that "if the breathalyzer detects any trace of alcohol, the vehicle will be seized and the driver subject to imprisonment".

> Ursula goes to a rave that lasts all night. To enjoy the 8 hours of electronic music playing nonstop, Ursula decides to make use of a drug called "ecstasy". This drug produces hallucinogenic effects (changes in perception) and stimulant effects (increased physical activity and insomnia). Ursula does not drink any alcohol at the party, only water. At sunrise, once the party is over, Ursula gets in her car to return home still under the influence of "ecstasy".

In the "no prompt" condition, we asked participants only if "Ursula violated the zero-tolerance policy". We then included either only the textual probe ("Ursula drove after ingesting a product containing alcohol") on the "semantic prompt" condition; only the moral probe ("Ursula is morally blameworthy for driving under the influence of 'ecstasy'") on the "moral prompt" condition; or both of them, on the "both prompts" condition. Each rule was presented to participants accompanied by either an under- (as above) or overinclusion case (e.g., describing someone who drives after using an alcohol-based mouthwash in the morning).

If the relevant difference introduced by Study 3 was to eliminate the priming effect of the semantic and moral probes, we should expect a significant difference between the "no prompts" and "both prompts" conditions. More precisely, participants under the "no prompts" condition should behave as they did in Study 3, ascribing more importance to purposes than to text, while those under the "both prompts" condition should behave as in Studies 1 and 2, privileging text. Moreover, the introduction of the "semantic prompt"' and "moral prompt" conditions allow us to evaluate whether the hypothesized shift in responses was caused by the introduction of one or the other question. In all conditions, participants agreeing more with the statement that the rule was violated in over- as opposed to underinclusion cases would show a preference for text, and the inverse pattern of results would reveal a preference for purpose

Figure 2 contains the key findings of our study, which relied on the responses of 364 participants (235 of which reported some legal training):

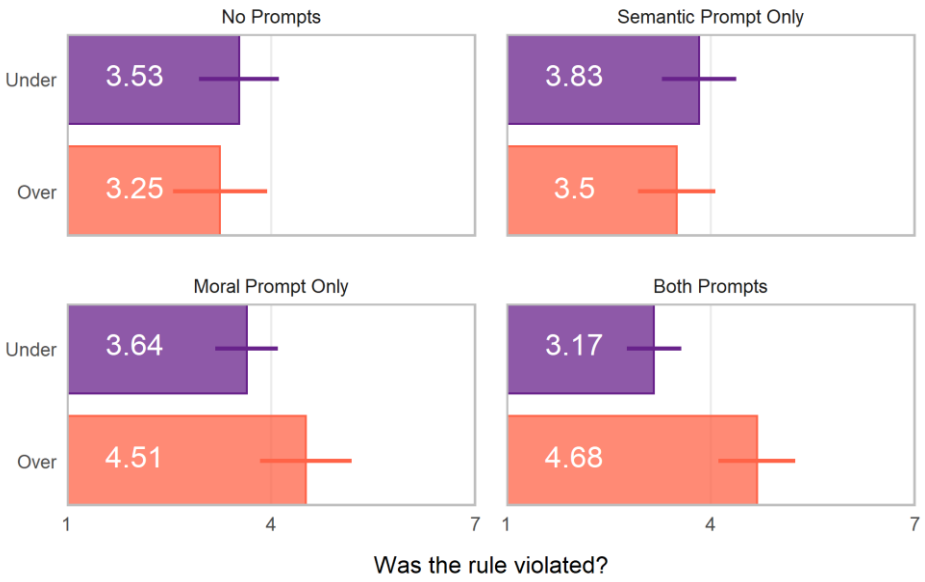

Figure 2 - In the "no prompts" condition, underinclusion cases are seen as rule violations more often than overinclusion cases, a picture that is inverted in the "both prompts" condition. Moreover, this seems to be driven mostly by the introduction of the moral prompt[62]. Struchiner, Hannikainen & Almeida (2020), p. 324.

The results confirmed our hypothesis: in the “both prompts” condition, but not in the “no prompts” condition, participants rated overinclusion cases to be rule violations more often than underinclusion cases. This difference seems to be more decisively driven by the introduction of the moral prompt than of the semantic prompt. Once again, the overall message stays the same if we restrict the analysis to subjects with legal training.

What can explain the puzzling finding that the questions we ask of participants affect what goes into their rule violation judgments? One candidate that we explored is the part-whole dynamic at play with the three questions. This topic was first explored by Schwarz, Strack & Mai (1991). In that paper, the authors asked people to rate their marital or dating life before assessing their overall life quality. Interestingly, the responses varied depending on conversational conditions manipulated by the researchers.

> If the questions were introduced outside of a joint evaluation context (i.e., if both questions were presented sequentially through no unifying context), participants showed an assimilation effect whereby there was a higher correlation between both answers than in a control condition.

[62] In Struchiner, Hannikainen & Almeida (2020), this same figure is identified as Figure 4.

> If, however, both questions were introduced by a shared prompt that made it salient that they pertained to one single conversational context, participants showed a contrast effect, whereby the correlation between questions was (non significantly) lower than in the control condition. The authors theorized that this is due to the Gricean maxim of non-redundancy: since participants have already expressed that they have a good (bad) romantic life, the remaining general question must be probing for something else that is unrelated to the first question (Struchiner, Hannikainen & Almeida, 2020, p. 320).

Something similar could explain the results of Study 4: since rule is a concept composed of two distinct elements, text and purpose, asking about only one of them and about the larger concept at the same time might lead to assimilation (in a non-conversational context) or contrast effects (in a conversational context). Consistent with the fact that questions were presented in a single unifying context, we observed a contrast effect when we introducing the moral probe: on that condition, participants answered the rule violation question in a more textual manner, possibly because of the applicability of the maxim of non-redundancy.

There are some problems with this interpretation. The first one is that, at first glance, the part-whole dynamic should predict that asking only the semantic probe alongside rule violation judgments should also lead to a contrast effect in the opposite direction. However, that is not what we observed. Even though the gap between under- and overinclusion cases is wider in the predicted direction in the "semantic probe" condition than in the "no probes" one, this difference is statistically insignificant.

Moreover, although part-whole relationships might account for the effects caused by the introduction of each question separately, they cannot fully account for the gap between the "no prompts" and the "both prompts" conditions. After all, in the "both prompts" conditions the presumption is that all subcomponents of the concept of rule (or at least all salient subcomponents) are also present.

An answer that we explored in "An experimental guide to vehicles in the park" is to revisit the conversational dynamics in light of the results. If we have one thing that is composed of two smaller, but equally important things, it is indeed the case that conversational norms should predict a contrast effect. However, if one of the components is much more important than the other, it might be more reasonable to expect effects of assimilation.

We tried to explain this idea by exploring an analogy with conversations about a meal (Struchiner, Hannikainen & Almeida, 2020, n. 19). If a meal is composed of equal amounts of risotto and beef, questions about one or the other followed by questions about the meal should result in contrast effects. Consider conversation 1:

> Part-question: How was the beef?
> Part-answer: It was OK, I guess…
> Whole-question: And how about the meal?
> Whole-answer: Oh, the meal was great!

In conversation 1, the whole-question seems to beg for a focus on the part that was not discussed yet. It seems reasonable to imagine that the correlation between the whole-answer and evaluations of how tasty that latent part was (the risotto) would be higher in that setting than if the whole-question was the only one asked. Thus, it really seems that, in a setting such as this, we should expect contrast effects when introducing part-questions before whole-questions in a single conversational context.

However, consider conversation 2, about a meal comprised of a large bowl of soup and three small pieces of bread:

> Part-question: How was the soup?
> Part-answer: It was great!
> Whole-question: And how about the meal?
> Whole-answer: Oh, the meal was only OK…

Conversation 2 does not sound natural, but it would be the exchange predicted by the account provided by Schwarz, Strack & Mai (1991). Here, it seems like assimilation should be expected: if the whole-answer was "It was awesome!", the exchange would make perfect sense. In "An experimental guide to vehicles in the park", we used this explanation to account for the pattern of results observed in the "both prompts" condition. But this explanation predicts an assimilation effect on the "semantic prompt" condition. However, that is not what we observed: the non-significant effect in that condition was actually in the direction of contrast, and not of assimilation.

In summary, a part-whole account that focuses solely on contrast effects provides a good explanation for the "no prompts" and "moral prompt" conditions, can accommodate the results in the "semantic prompt" condition, but fails to

explain the "both prompts condition". On the other hand, a part-whole account that posits that the textual element is so prevalent that we should expect an assimilation effect gives a good explanation for what goes on under the "no prompts", "moral prompt" and "both prompts" conditions, but is directly contradicted by the "semantic prompt" condition. Hence, part-whole dynamics alone offer a messy explanation for the results. Each of the two versions of part-whole explanations considered above leaves one of the conditions unexplained.

A slightly different way to conceptualize the results combines elements of the part-whole explanation with an explanation based on the differences between analytical and holistic styles of thinking. Perhaps, the introduction of part-questions predisposes participants to engage in a kind of inquiry that is qualitatively different from the one they would otherwise pursue. While a simple "did John break the rule" kind of question engages people in all-things-considered intuitive reasoning, a richer probe explicitly mentioning text and purpose leads them to reflect about each element before reaching an overall judgment[63]. Under that account, the effects caused by the introduction of each prompt could be explained mostly by part-whole dynamics, but their combined effect would amount to something greater, as is indeed observed in the (statistically insignificant) gap between the "moral prompt" and "both probes" conditions.

Positing that analytical thinkers would prefer text over purpose, but that holistic interpreters would show a preference for purpose has the added bonus of accounting for the differences between real life textualists and purposivists. For

[63] Recent work (Kneer, Colaço, Alexander & Machery, in prep.) investigates the "reflection defense", the claim that experimental philosophy is irrelevant for the investigation of philosophical claims because it fails to probe intuitions that are the result of careful and reflective thought. The authors show that a host of strategies fail in making lay people's responses follow the patterns predicted by philosophers in Gettier cases and other famous thought experiments. They tried forcing a delay between questions and answers, financially incentivizing participants to get the right answer to philosophical puzzles and asking participants to solve mathematical problems before answering the philosophical questions, but none of these manipulations mattered to a significant extent. They also used well validated measures of cognitive style to show that interpersonal differences in reflectiveness and intuitiveness could not explain the results. At first glance, this finding might be read as casting doubt on the explanation for Struchiner, Hannikainen & Almeida (2020) that relies on a difference between analytical and holistic thinking. However, my claim is not that one style of thought is more careful, reflective or profound than the other. The mention to "analytic" here is also not in reference to the a priori vs. a posteriori distinction, but merely to the strategies of either decomposing a concept into constituent elements versus taking the entire concept into consideration without taking these decomposing steps. Hence, the results of Kneer and colleagues tell us little about what to expect of this explanation here, as they did not test for differences between this specific dimension of the analytic-holistic divide.

instance, Hart was a prototypical *analytical* philosopher, while Fuller was avowedly suspect of this kind of reasoning, taking himself to be engaged in a different sort of enterprise[64]:

> The term 'analytical' is also apt in conveying an intellectual mood that finds more satisfaction in taking things apart than in seeing how they fit together; there is, indeed, little interest among analytical positivists in discerning the elements of tacit interrelatedness that infuse – though always somewhat imperfectly – what we call, by no accident, a legal *system* (Fuller, 1969, p. 191).

Hence, explaining the results of Study 4 partly with reference to the analytic-holistic divide has the theoretical virtue of accounting for the way the debaters themselves understood their differences. The analytic-holistic distinction by itself, however, cannot fully account for the difference caused by the introduction of the "moral prompt" alone. Thus, there is a need to complement it with some elements of the "part-whole dynamics" explanation. For instance, we could argue that purpose is the least salient of the elements composing the concept of rule. Thus, making it explicit could be sufficient to make most of us approach the rule violation question in an analytical frame of mind. As this relies on the part-whole dynamics explanation to account for one of the conditions, this would be a hybrid approach that combines elements from this explanation with the analytic-holistic distinction.

What are the implications of Study 4 for the Hart-Fuller debate? The first thing to note is that the simple deflationary answer to Study 3 ("purpose matters only in the penumbra") is no longer available. Something (maybe part-whole dynamics, maybe analytic vs. holistic thinking, maybe some combination of both) about the questions being asked matters to the relative importance of text and purpose. Hence, TP and WP are each true under different settings: when we consider rule violation judgments holistically, without taking into consideration how the situation at hand relates to text and purpose separately, WP is true, but when the distinct roles played by text and purpose are made explicit (which might already happen to a large extend when purpose alone is made salient), TP obtains.

Each of the explanations considered above presupposes that rule is a composite concept that has at least two components, text and purpose. Moreover,

[64] Barak (2005) also makes repeated claims that purposivism is holistic.

one of the components is descriptive (text), while the other is normative (purpose). This makes the concept of rule a dual-character concept, like the concepts of reasonableness (Tobia, 2018), normality (Bear & Knobe, 2016), and happiness (Phillips et al., 2017), among others. I will call this explanation, or family of explanations, the "dual nature of rule" explanation (DNR):

> Dual nature of rule (DNR): the concept of rule is composed of a descriptive (text) and a normative (purpose) component.

3.1.5 Evaluating the evidence

If DNR provides the best explanation for Studies 1-4, what follows with regards to the four theses? In each and every Study, I have argued that the evidence supported CSM against SP. The weak and relatively easy-to-find-support-for Hartian thesis trumped the strong and nearly impossible to confirm Fullerian thesis. One easy way to close the debate is to circumscribe it to those broad conceptual claims and declare Hart the winner. But this might be too easy. SP might just have been a hyperbolic way of stating Fuller's deeper insight that we would miss something very important in law if we neglected the role of purpose[65]. As Hart noted in his review of Fuller's "The Morality of Law", "[t]he author has all his life been in love with the notion of purpose" (1983, p. 363). Thus, even if purpose isn't *always* a *necessary* condition for rule violation (SP), finding that it is often important[66] and sometimes at least as prevalent as text (WP) surely vindicates Fuller to some extent.

There is, however, another way to interpret the results that would lead to a completely different theoretical picture. This alternative reading discounts the influence of purpose as the result of an indirect speech act. The idea finds

[65] See the discussion in Rundle (2012), p. 6, claiming that Fuller's claims were ridden with "numerous instances of hyperbole". Unfortunately, the book does not go on to discuss all of those hyperboles in detail, but I would be tempted to say that SP might have been a fruit of one them.

[66] Even where TP holds (Studies 1 and 2), the effects of purpose were still significant. One way to see this is to think that overinclusion cases were not judged to be rule violations all the time, with sizeable minorities reporting that the rule was not violated in those circumstances. In other words, there was an important difference between overinclusion and core cases. In the latter, where text and purpose agree, almost everyone reports that the rule was violated, in contrast to a more contested majority in the former. The same is true about underinclusion cases: even if only a minority deemed such cases to be violations of the rule, that minority was sizeable enough as to distinguish underinclusion from off-topic cases.

expression in work done by John Turri and Peter Blouw about blameless violation of rules.

3.2
The challenge posed by indirect speech acts

Turri & Blouw (2015) take as a starting point the work of Mark Alicke. Alicke (1998) found that people were willing to assign more responsibility when a protagonist acted for socially unacceptable reasons than when it performed the exact same act for acceptable reasons. For instance, in one study, he presented people with the following stimuli: "John was driving over the speed limit (about 40 mph in a 30 mph zone) in order to get home in time to" either "hide an anniversary present for his parents" or "hide a vial of cocaine". Subsequently, something else (for instance, some oil that spilled in the road) contributed to cause an accident that left another driver with "multiple lacerations, a broken collar bone, and a fractured arm" (p. 369). People in the "cocaine" condition considered John to be more of a cause of the accident, and assigned him more responsibility for the bad outcome. In later work (Alicke, 2000), the author explained this effect as being caused by "blame validation processing", which "refers to observers' proclivity to favor blame versus nonblame explanations for harmful events and to de-emphasize mitigating circumstances" (p. 568). According to Alicke, this leads people to posit that those who are to blame probably knew and had more control over the outcome than those who are not.

Turri & Blouw (2015) set out to explore whether blame validation is "only half the story". In other words "Might people also engage in *excuse validation* too? That is might people also interpret the facts so that it validates their desire to excuse?" (p. 616). At first, this discussion might seem completely alien to the debate about rules. But the way Turri & Blouw opted to test their hypothesis makes use of rules in a way that is very relevant for our concerns.

3.2.1
Study 5: excuse validation (Turri & Blouw, 2015, Experiment 2)

Study 5 consisted of a set of vignettes where:

> the protagonist has good evidence to think that they are following the relevant rule and correctly engaging in the activity, but something outside of their awareness foils them and they end up breaking the rule anyway. For example, in the Driving story [scenario], Doreen just had her car serviced and is driving home from the mechanic's shop. She wants to get home without unnecessary delay, but she does not want to break any traffic laws. The traffic law sets the speed limit at 55 miles per hour, so she looks down to see how fast she is going. The speedometer says that she is going 55 miles per hour. But the speedometer malfunctions and underestimates the car's speed. Doreen is not aware of this and is actually driving 60 miles per hour (p. 618).

Participants' were then asked a series of questions including "Should Doreen be criticized for driving that speed?" and "did Doreen break the rule?". The way the authors phrase their hypothesis is very instructive about the concept of rule they presuppose:

> The protagonist in the story breaks a relevant and salient rule. […] Thus one might reasonably expect, at the very least, a very strong majority of participants to agree that [the protagonist broke the rule] […]. For example, in the Driving story Doreen is driving 60 mph in a 55 mph zone. It follows unmistakably that she is breaking the speeding law. […]
> However, the protagonist is justified in thinking that they're following the rule and they end up breaking the rule in light of misleading evidence. […]
> Cases like this provide an excellent test for whether excuse validation actually occurs. If excuse validation does occur, then many participants will deny that the protagonist's behaviour [breaks the rules]. But if excuse validation doesn't occur, then it's unlikely that many people will deny that the protagonists [broke the rules]. Instead, we'd expect almost everyone to agree that [they broke the rules], which is the objectively correct answer (p. 619)[67].

Throughout the setup, the meaning of the rule is equated with its text (presumably something like "no driving over 55 miles per hour"), with no consideration about the requirements of the rule's purpose. Hence, the authors find it very surprising when only 45% of people in the Driving story stated that Doreen broke the rule. Obviously, Turri & Blouw were not seeking to establish anything

[67] The only difference between Experiments 1 and 2 in Turri & Blouw (2015) is that the test question in Experiment 1 is formulated in terms of whether the protagonist was in any sense incorrect, while Experiment 2 formulates it explicitly in terms of rule violation. However, the Experiments' share the same setup. To facilitate the exposition, I have changed the way the reference question is worded in the excerpt above between brackets.

about the concept of rule, but it is interesting to consider what a Schauerian analysis of their Driving scenario has to say and whether it might help explain their puzzling results.

The first step in such an analysis is to ask what is the purpose of the rule that establishes the speed limit. I think that the answer here is quite straightforward. Even if this is not explicitly stated in their story, most people would surely think that the rule's purpose is "to avoid accidents", or something along those lines[68]. Assuming that this is a well calibrated rule, it is generally the case that driving over 55 mph increases the likelihood of accidents significantly when compared with driving under that speed limit, at, say 50 mph. Thus, at first glance, the right classification for Doreen's case is as a core case of rule violation: after all, she is both violating the rule's text by driving at 60 mph and the rule's purpose by increasing the likelihood of an accident.

But that is not entirely clear. People might believe that the rule is not well calibrated (thinking something like "actually, 55 mph is way too low, only at 65 mph the risk of accidents becomes unacceptably high"), or that Doreen would be sensitive to the malfunctioning if she was truly increasing the probability of an accident ("if she wasn't aware that she was in the wrong – because she wasn't driving any faster than anyone she could see, for instance – that's evidence that she was probably right"). In other words, the experimental design (which was not intended to test the concept of rule) was not construed in a way capable of completely ruling out the sort of moral evaluation that is introduced by purpose in rule violation judgments. So, one (tentative) way to make sense of Turri & Blouw's results is to paint them as more evidence of the importance of purposes in rule violation judgments. That, however, is not the route they take.

3.2.2
Study 6: indirect speech acts (Turri & Blouw, 2015, Experiment 4)

The authors introduce Study 6 in the following way:

[68] I have no quantitative evidence for this, but every single time I have used speed limits as an example in class, I have asked students what was the purpose of the speed limit, and in every single time they answered that the purpose was to avoid accidents.

> Why do people engage in excuse validation? […] Saying that someone "broke a rule" […] is often a way of *indirectly* criticizing or blaming them […]. But participants don't want to (appear to) unfairly criticize the protagonist. So instead of agreeing that the protagonist "broke the rules", which would be true but misleading, many participants disagree. In other words, many participants say that the protagonist did not break the rules because it's the only available option that won't implicate them in unwarranted criticism (p. 623).

To test this, they used stories with the same structure as Doreen's, but changed the questions asked. Now, participants were required to choose one of the following two alternatives: "the protagonist unintentionally broke the rules" and "the protagonist did not break the rules". Under that set up, 91% of people answered that the protagonist unintentionally broke the rule. Further studies in the same paper showed that this is not merely a feature of this specific forced choice design.

3.3
Reevaluating the evidence

Study 6 provides evidence for what I will call the "indirect speech act explanation" (ISA):

> Indirect speech act (ISA): rule violation judgments that deviate from the text are entirely the result of indirect speech acts regarding the blameworthiness of the protagonist.

ISA negates DNR insofar as it presupposes that a rule is indistinguishable from its text. Under that account, rules don't have a dual nature. They have a singular textual nature. Every deviation from the rule's text must hence be explained by some extrinsic, distorting factor. Here, ISA adds that the correct explanation is the existence of an indirect speech act.

Other people that share Turri & Blouw's assumption that rules have a single nature, but disagree that speech acts have anything to do with it might come up with other explanations for the deviations. One could say that participants are confused by the wording of the question, are answering to some implicit demand from the experimenter, are motivated to give a certain response, or point to a host of other factors that might interfere with people's capacity to apply the (under that view)

textual concept of rule[69]. In any event, as far as denials of DNR go, ISA seems to be the better candidate, as there is direct evidence for its truth in Study 6.

How well does ISA fare as an explanation for Studies1-4? As formulated, very well. Although ISA was advanced to explain a phenomenon that happened in putatively core cases, phrasing it in terms of blameworthiness makes it suitable to deal with both over and underinclusion cases.

In overinclusion cases, the existence of a significant minority that says the rule was not violated could be explained by an indirect speech act built into statements such as "John broke the rule". That assertion might entail something like "John is to blame". But John seems blameless in overinclusion cases. Thus, people excuse him, making it so that less people agree with the rule violation statement in overinclusion cases when compared to core cases. In underinclusion cases, the statement "John did not break the rule", by its turn, might entail something like "John is not to blame". But here, John is clearly to blame (i.e., for taking his pet bear into a crowded restaurant that enforces the "no dogs" rule). As a result, blame validation (Alicke, 1998; 2000) kicks in and leads people to say that the rule was violated more often than they would do in an off-topic case.

This makes it clear that ISA might account for the results observed in Studies 1-3. According to it, what we (Struchiner, Hannikainen & Almeida, 2020) labeled as an effect of purpose was actually just the result of indirect speech acts about blameworthiness; since our questions that were putatively about purpose actually tracked blameworthiness (as is explicitly the case in Studies 3 and 4), both ISA and DNR would predict the same pattern.

What makes ISA especially plausible, however, is how well it can account for the results of Study 4. If rule violation judgments are entirely textual, but our results are confounded by an indirect speech act regarding blame, giving participants the option to say something like "John broke the rule, but is not to blame", like Turri & Blouw did in Study 6, should lead to increasingly textualist responses. And, lo and behold, that is exactly what Study 4 shows: when we

[69] Another way to put that is in Knobe's (2010) terms: DNR views the influence of moral factors as a feature of our concept of rule that is built in to our competence to use it, while single nature views will tend to discount the effects of moral factors as a bias.

introduced the moral, but not the textual prompt, participants started to see overinclusion cases more (and underinclusion cases less) as rule violation cases. Arguably, this is a simpler and hence better explanation for Study 4 than DNR's part-whole/analytic vs. holistic explanations.

But there are also some things that ISA cannot explain in Studies 1-4. In Study 6, the introduction of the blameless violation option meant that over 90% of the participants began to apply the concept of rule in a textual manner. In Studies 1, 2, and 4, on many conditions, participants were asked about rule violation and blame concurrently (under the guise of purposive violation in Studies 1 and 2, and moral blameworthiness for such violation in Study 4). In none of those Studies, however, such a large proportion of participants followed the pattern predicted by a purely textual concept of rule. Why is it so? A proponent of DNR might point to the fact that Study 6, just as Study 5 before it, deals with something that is not clearly a clash between text and purpose. Perhaps on those occasions there is truly an indirect speech act distorting judgments, but this might still not be the correct explanation for what goes on in under- and overinclusion cases.

The upshot is that Studies 1-6 are simply not enough to adjudicate between ISA and DNR. Both explanations have virtues and problems. DNR's main problems are, on the one hand, its convoluted way to account for Study 4, and, on the other, the results of Study 6. Its main virtue is the ability to account for the pervasive statistical significance of purpose, even on conditions where participants had a way to express the indirect speech act that might be conveyed by rule violation judgments. ISA, by its turn, cannot fully account for that, and this is the main detraction for this explanation. However, it offers a superior explanation for the results of Study 4.

3.4
Rule-identity judgments

In order to test the relative merits of ISA and DNR, we need a setting where relevant indirect speech acts are unlikely to occur. If there is a situation where we can probe people's concept of rule without having them judge anyone in particular, we might see whether purpose still interferes. If it does, that speaks in favor of

DNR. If it doesn't, then ISA is more likely true. That is precisely the test we conducted in a paper currently under review (Almeida, Struchiner & Hannikainen, under review).

We turned to a body of research about personal identity for inspiration on how to do away with rule violation judgments. This literature shows that morality shapes folk judgments of personal identity[70]. The experimental paradigm involves describing someone that undergoes some radical change that is of moral significance and then asking if the person after the change is the same as the person before the change. Tobia (2015), for instance, found that whenever people undergo moral improvement, changing from bad to good behavior, most participants say that the person is still the same. In contrast, when someone good undergoes change to become morally bad, people agree with the idea that "it seems like one person died […] and it is really a different person entirely that exists after [the change] […]" (p. 398).

The details and interpretation of these results are irrelevant here, but the methods employed are not. ISA offers a good explanation for Studies 1-6 because they all involve *rule violation* judgments. But if we are interested in the nature of the concept of rule, as DNR is, we might investigate the roles that text and purpose play on *rule identity* judgments.

According to ISA, rules are nothing but their texts. Hence, whenever a rule's text change, its identity should also change. However, if text stays the same and there is a change in the rule's purpose, people should still think that they are dealing with one and the same rule. In contrast, DNR predicts that both changes in text and purpose will influence rule identity judgments.

3.4.1 Study 7

We came up with an experiment that manipulated how a rule evolved over time to probe people's intuitions about rule identity. Below, I reproduce the text for all the different vignettes under the "no vehicles in the park" rule, to keep in the

[70] See Newman, Bloom & Knobe (2014) and Tobia (2015) for two examples of this line of inquiry. Strohminger, Knobe & Newman (2017) offer a recent overview of this literature.

theme of the Hart-Fuller debate. The experiment also included other 3 rules: “no dogs in the restaurant”, “no shoes in the apartment”, and “no driving over 60 km/h”. All conditions were introduced with the following prompt:

> **Shared prompt**
>
> John, mayor of Mirandopolis, aiming to cut down the noise that hindered the local park's tranquility, created a rule prohibiting vehicles in the park. To communicate this rule, signs were posted at all entrances saying "No vehicles in the park".
>
> Two years later, John's mandate is over and Peter is elected the new mayor.

After that, participants saw one out of the following four conditions:

> **Nothing changed:**
>
> Peter shares with John the goal of cutting down the noise and keeps all signs up without any changes.
>
> **Only text changed:**
>
> Peter, while sharing with John the goal of cutting down the noise, believes that the rule's text is not clear enough and changes all signs for signs that read "No motor vehicles in the park".
>
> **Only purpose changed:**
>
> Peter keeps all signs, but his goal is different than John's. Peter believes that vehicles are forbidden not because of the noise they make, but because of the danger they bring to park goers physical integrity. Soon after taking office, he announces his decision to the public saying: "I don't care at all about cutting down the noise; all I care about is safeguarding park goers physical integrity".
>
> **Both changed:**
>
> Peter believes that vehicles are forbidden not because of the noise they make, but because of the danger they bring to park goers physical integrity. Soon after taking office, he announces his decision to the public saying: "I don't care at all about cutting down the noise; all I care about is keeping park goers physical integrity".
>
> Not only that, but he also believes that the text is not clear enough and changes all signs for signs that read "No motor vehicles in the park"[71] (Almeida, Struchiner & Hannikainen, under review).

[71] A translation of the complete set of stimuli is available at https://osf.io/xtswd/?view_only=9cf87d9a516e4d03b23809b6f5de759d.

302 participants, including 163 reporting some legal training[72], completed the survey. Each participant received the same condition under all rules on a random order. For each case, they answered how much did they agree on a 7-point scale ranging from "completely disagree" (1) to "completely agree" (7) with the statement that "the rule created by John is the same rule used by Peter".

The results showed a significant difference between the conditions where nothing changed and where only purpose changed, indicating that purposes affect rule identity judgments. However, the influence of text was decisively greater than that of purpose. In fact, mean ratings for both the nothing changed (6.37) and only purpose changed (5.44) conditions fell significantly above the midpoint of the scale (4), indicating that most people felt that the rule was mostly – but, crucially, not completely – the same after purposive change. This is consistent with DNR: saying that the concept of rule is comprised of textual and purposive elements does not entail that these elements are equally important. ISA, however, fails to provide an adequate explanation for the data, as it is much harder to argue that the answers to the rule identity assessment entail any judgments of blame.

But post-hoc analyses complicate the picture. When we consider the data separately for lawyers and lay people, we see two distinct pattern of results, depicted in Figure 3:

[72] See details about the legal expertise question in footnote 55, supra.

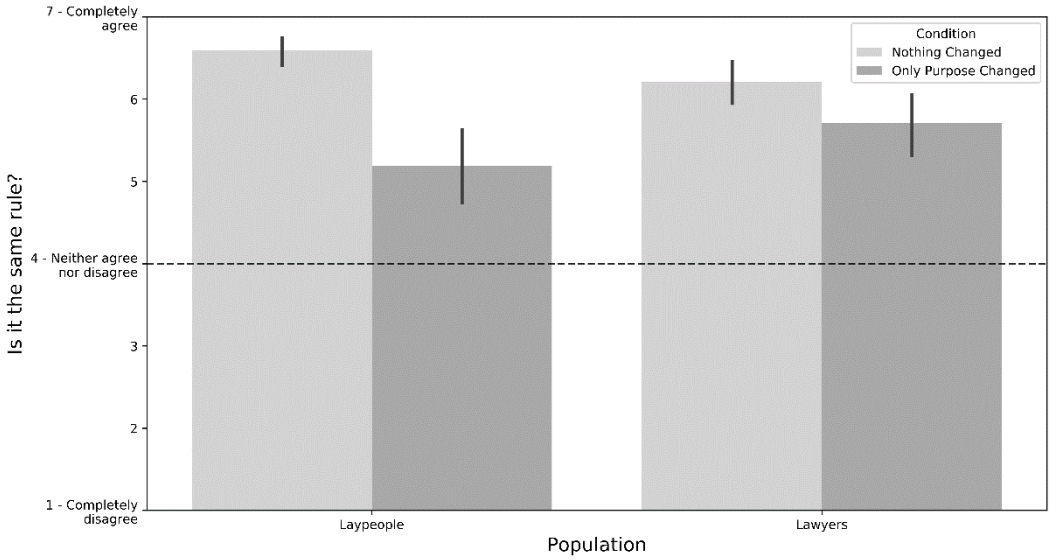

Figure 3 - Rule identity judgments differ depending on participants' legal expertise. Those without legal training show a significant effect of purposive change, while lawyers do not.

This difference between populations was unexpected. In Studies 1-4, all relevant effects remained significant when we excluded laypeople. Hence, the working assumption of Struchiner, Hannikainen & Almeida (2020) was that both laypeople and lawyers shared a single concept of rule that was composed of the same elements. These results, however, might suggest a different picture.

Specifically, if we take Study 7 seriously as a test between ISA and DNR, it suggests that ISA provides the best explanation for lawyers, while DNR best captures laypeople's intuitions. Revisiting the implications of each explanation for the theses of the Hart-Fuller debate, we can recall that ISA provided an account that fully vindicated Hart's theses, while DNR made room for Fuller's deepest insights. Therefore, assuming that Study 7 reveals the right picture, it could be said that Hart was completely right about the concept of rule employed by lawyers, while Fuller's focus on purposes capture something true about the way laypeople, but not professional lawyers, deal with rules. Perhaps not by coincidence, Hart was predominantly concerned with the intuitions of officials, while Fuller gave great importance to the perspective of citizens.

Why do lawyers and laypeople diverge? Previous work on systematic differences between those populations (Tobia, 2020b) suggests that lawyers sometimes use the same word (e.g., intentionality) to designate a slightly different

concept than that ordinarily employed by laypeople. These concepts might be acquired in at least two different ways: training or experience. Maybe lawyers employ some concepts in a way that is different than laypeople because they were taught so in law school. Alternatively, perhaps the driving force isn't academic learning, but experience: applying these special concepts day in, day out, slowly but surely makes lawyers develop different intuitions about them. If the populational differences are mostly driven by training, law students should behave very similarly to experienced judges and both groups should diverge from laypeople. On the other hand, experienced public officials that deal with rules on a daily basis, but never went to law school, should have similar intuitions to the general population. If, instead, the differences are caused by experience, experienced public officials should be more like judges than the general public, and inexperienced law students should behave as laypeople.

Which one is it? Kahan et al. (2016) investigated whether professional judges, lawyers, law students, and members of the general public were influenced by their political views when judging rule violations. They found that members of the general public tended to make decisions consistent with their political outlook, while lawyers and judges did not. Law students, on the other hand, were still affected by political bias, but to a lesser extent than the general population. Kahan and colleagues remark, based on this set of results, that "The process of acquiring this species of professional judgment obviously does not end in law school. But our study suggests that it certainly begins there" (p. 414). One way of reading these results suggests that both experience and training contribute to legal expertise. As such, although we are using legal training to classify participants as laypeople or lawyers, future research should also look at their experience to check whether it includes the creation and application of rules.

3.5
Recognitional communities

In the introduction, I have argued that Hart's position in "The Concept of Law" is that only the beliefs and attitudes of officials towards legal rules are necessary for the existence of a legal system. In contrast, Kristen Rundle (2012) argues that one of the main points of contention between Hart and Fuller was

precisely that Fuller sought to develop a theory of law that also took into account the perspective of the addressees of rules[73]. While Fuller was never explicit with regards to this issue, I find Rundle's argument convincing. She reviewed extensive evidence, including Fuller's notes and correspondence, and grounds her interpretation on a holistic assessment of this material. Assuming that both of these interpretations are correct, can this disagreement about recognitional communities explain the gaps between CSM and SP, TP and WP? I believe it can.

Recall that in the introduction I have argued that the examples proposed by each side on the 1958 exchange were appeals to the intuitions whereby each philosopher invited us to think whether ourselves, and perhaps, a given subset of the general population, shared his own intuitions about certain cases. Theory building then followed in one or the other direction depending on which intuitions were assumed. Back in the introduction, I ventured that the intuitions underlying each appeal were in conflict, such that Hart appealed to a textualist intuition, while Fuller appealed to a purposivist intuition.

An alternative to viewing the appeals to intuitions as in conflict, is to construe them as having *different targets*. Perhaps, Hart was appealing to the intuitions of *officials*, while Fuller appealed to the intuitions of *citizens*. Now, if it turns out that officials and citizens have different intuitions altogether, then there is no contradiction in the original pair of appeals. After all, it is perfectly possible that officials share Hart's textualism, while laypeople share Fuller's purposivism. Moreover, that seems to be the message coming out of the data: lawyers – a (non-random) sample of officialdom (more on this in a moment) – behave like Hart would have predicted, and ordinary citizens showed the pattern of results of (weak) purposivism.

The very fact that lawyers and laypeople employ the concept of rule differently – although not entirely new, see Schauer (2009) – might thus hold the key to understanding this longstanding debate in legal philosophy. Potentially,

[73] It would be wrong to say that Hart doesn't take this point of view into account. His legal theory is prominently about explaining the *internal point of view* (see Shapiro, 2006), and in healthy legal systems citizens also adopt that stance (as argued in the introduction). The most charitable way to read Rundle's and presumably Fuller's point is about taking issue with the narrower question of recognitional community.

clarity with regards to recognitional communities might prove fruitful beyond the Hart-Fuller debate. Adler (2006) argues that a similar divide in recognitional communities separates Hart and Dworkin[74], as well as several other groups of scholars. Making the target of an appeal to intuition explicit might thus show how some apparently contradictory accounts of law might turn out to be actually complementary.

However, this newfound explanatory power of recognitional communities relies on two assumptions that need to be further spelled out and tested: 1) there is an important relationship between the rule of recognition broadly construed and the concept of rule employed by the recognitional community; 2) we can speak interchangeably of lawyers and officials. Since they are important for the claim that recognitional communities matter, I sketch a brief defense of each assumption in what follows.

3.5.1
Rules and the rule of recognition

References to the rule of recognition often involve an ambiguity between process and results. First, we might refer to the rule of recognition as a process: the process according to which a determinate community develops the attitude of using certain rules as a means to guide their own conducts and as a standard by which to judge other peoples' behaviors[75]. In other words, when discussing the rule of recognition, we are sometimes discussing how people come to accept a system of rules. Second, we might refer to the rule of recognition as the result of this process: the actual set of criteria that defines what it takes for a rule to be a part of a particular legal system.

Recognitional communities that employ the concept of rule differently during the *process* of recognition might refer to the same sources (i.e., constitutions,

[74] "On the Dworkinian account, U.S. law at present is identified by a constructed rule of recognition that integrates considerations of straight moral justifiability with the preinterpretive understanding of U.S. law held at present by U.S. citizens: a shared, contemporaneous citizen understanding that, in turn, will identify certain past events as legal decisions (as "acts of Congress", "Supreme Court decisions", etc. My deep popular constitutionalist reading of Dworkin's Law's Empire is surely contestable […]. [But] that inclusive view of the U.S. recognitional community is a natural inference from Dworkin's argument for why all U.S. citizens have a moral obligation to obey U.S. law" (p. 742).

[75] Hartian terminology takes adherence to this pair of attitudes to define the "internal point of view" (see Shapiro, 2006).

statutes, offices, institutions…), but end up guiding and evaluating behavior by very different standards[76]. For instance, one group of people who adopts a very purposive concept of rule might end up recognizing a set of rules that is different than the one recognized by a group of staunch textualists. In other words, the way people interpret the sources during the process of recognition shapes the content of the recognized system of legal rules. Since different groups of people interpret rules differently, our account of what it takes for something to count as law (the rule of recognition as a result) will depend on the recognitional community we focus on.

Study 7 suggests that lawyers always identify the scope of rules with reference solely to text. Under the view suggested by the results of this study, moral considerations might play a role in legal decision-making, but they do not play any role in the way lawyers determine the content of rules. Focusing on this recognitional community would lead to the conclusion that rules are taken to apply even to those cases where the results they recommend are morally defective. Hence, the best theory to describe what counts as law to this group would be positivistic in nature. In contrast, laypeople do take purpose into account when making even rule identity judgments. It is thus plausible that the best theory to describe what counts as law to laypeople may include a moral test.

Hart's recognitional community is traditionally construed as including all legal officials. Are lawyers officials in the Hartian sense? This is the issue I will turn to on the next subsection. At this moment, it is enough to note that if we are warranted in treating lawyers and officials as similar categories, then Hart would be fully vindicated by the data. After all, lawyers behaved just as predicted by TP.

What about Fuller? If his focus was really on citizens as a whole, he might just as well also be descriptively right, for the data reviewed above suggest that ordinary citizens do take purposes into account when identifying the contents of a rule (WP). Hence, the content recognized by this community will presumably be a matter answerable to the moral sensibilities of the community (if purposes are really moral in nature, a view we discuss in chapter five). Thus, a salient feature of those

[76] About the source-norm distinction, see Shecaira (2015).

legal communities (assuming that citizens systematically moralize purposes) will be the congruence between law and morality.

That way to read the evidence suggests that Fuller and Hart were talking past each other in a way that hasn't been fully contemplated by existing scholarship. Even though both were really concerned with making claims about (1) the psychology of rule-based judgments, (2) the insights this psychology offered about the concept of law, and (3) their connections with morality, they were talking about the psychology of different groups.

How should we resolve this issue? Adler (2006) suggests that perhaps the right stance is to recognize one specific form of legal pluralism. Within a given territory, there will be as many legal systems as there are recognitional communities, even if they share many institutions. I take this position to be too extreme. There seems to be something that makes Brazilian law one single discernible system, despite the different attitudes that subsets of the Brazilian population might have towards rules. But there is some truth to the idea that recognitional communities matter. Thus, Brazilian law will likely have different contents depending on whether only officials or if the whole population internalize the legal rules[77]. Likewise, variations in the intuitions of the population at large and the relationships between officialdom and citizenry in different countries will likely lead to variations in the content and shape of legal systems across cultures, a subject I will return to in chapter five.

3.5.2
Lawyers and officials

Even though many officials are not lawyers, people with legal training are disproportionately represented in Brazilian officialdom. That they represent the totality of judges and a majority of judicial officials is a given, but the influence of lawyers extends well into the two other branches.

[77] Much more should and could be said about this, but I believe that this is sufficient to show how some of the jurisprudential differences between Hart and Fuller (and possibly other divides) might be caused by differences in the recognitional communities those philosophers had in mind.

Starting with the legislative branch: Sadek & Dantas (2000) report how lawyers occupied an overwhelming majority of seats in both the Brazilian house of representatives (64.8% in 1886) and the senate (72% in 1871) from the late 19th to the early 20th century. Although much less dominant than they once were, lawyers are still the largest professional group represented in Congress: according to data provided by the Brazilian house of representatives[78], at least 87 (17%) out of the 513 representatives elected in 2018[79] were lawyers. In contrast, there are currently slightly over a million registered lawyers in Brazil[80], approximately half percent of the country's current population (circa 212 million inhabitants in 2020)[81]. Lawyers also dominate elective positions in the Executive branch: a single law school in Brazil (USP) educated 13 different presidents. But lawyers' influence extend over officialdom in other, more subtle ways. Almeida (2010, chapter 6) reports that political parties, congressional cabinets, ministries, and other official positions in both the Legislative and Executive branches often employ large numbers of lawyers to shape public policy and political strategies.

So, at least in Brazil, where Studies 1, 2, 3, 4, and 7 were conducted, there is a non-trivial overlap between lawyers and officials. Is this a peculiarity of the Brazilian system? Although a comparative assessment is well beyond the scope of this dissertation a cursory glance over U.S. politics suggests otherwise. For instance, 26 out of the 45 presidents of the U.S. were lawyers[82]. They are similarly overrepresented in the U.S. congress: "While comprising a mere 0.4% of the voting-age population, lawyers accounted for 39% of seats in the House and 56% of seats in the Senate in the 115th Congress. The overrepresentation of lawyers vastly exceeds even that of millionaires" (Bonica, 2020, p. 253). Given these parallels, it

[78] https://www.camara.leg.br/internet/agencia/infograficos-html5/composicaocamara2019/index.html#text3

[79] 78 representatives are labeled as lawyers, while 9 are reported to have graduated from law school. This distinction is important, because the label of "lawyer" is reserved to those admitted to the bar association. Besides those, a number of other positions (e.g., high ranking police officer) are very often occupied by people with legal training. As such, the reported proportion almost surely underestimate the proportion of representatives that studied law.

[80] https://www.oab.org.br/institucionalconselhofederal/quadroadvogados

[81] https://www.ibge.gov.br/apps/populacao/projecao/box_popclock.php

[82] https://www.americanbar.org/groups/bar_services/publications/bar_leader/2009_10/january_february/presidential/

is not unreasonable to assume that the abundance of lawyers in auxiliary positions to officials is also something that occurs in the U.S.

Hence, at the very least, knowing that lawyers differ from laypeople in their intuitions about rules should increase our credence that officials will also differ from the general population in similar ways. What's more, if experience, instead of education, is the primary reason why lawyers and laypeople behave differently, we should expect officials who never went to law school to behave much more like lawyers than laypeople. Nonetheless, further research should test directly whether a sample of broadly construed officials would apply the concept of rule in the same way as lawyers do. Until then, it is plausible to speculate that the samples of legally trained subjects surveyed in Studies 1, 2, 3, 4, and 7 are indicative of the likely intuitions of officials at large.

3.6
Summarizing the evidence

In this chapter, I reported the results of seven different studies originally presented in three recent papers. For each study, I considered the implications of the results for the theses of the Hart-Fuller debate. Overall, CSM consistently trumped SP, but TP and WP traded blows. Study 4 suggests that whenever people are prompted to make rule violation judgments in isolation, without reference to any of the components of the concept of rule, WP prevails. However, under circumstances that explicitly draw attention to the distinction between text and purpose, TP holds. Why is that so? DNR proposes that rule is a dual nature concept consisting of a normative component (a rule's purpose) and a descriptive component (its text). ISA, supported by Studies 5 and 6, explains away the effects of purpose as stemming from our unwillingness to assert to an indirect speech act entailed by statements such as "Joan broke the rule". Studies 1-6 are compatible with both explanations, but Study 7 was designed to bring them apart. By asking about rule identity, instead of rule violation, indirect speech acts of blame seem to drop out of the picture. Surprisingly, in this setting laypeople incorporate purpose into the concept of rule, but lawyers do not. Thus, one reasonable interpretation is that DNR is true for laypeople, while ISA is true for lawyers. This, in turn, might help recontextualize the conflict in Hart and Fuller's appeal to intuitions. Perhaps,

the authors had different recognitional communities in mind and were thus talking past each other in a subtle way. Under that reading, both are vindicated by the data and the broader issues of the debate should be recast in terms of the appropriate recognitional community.

Hart was explicitly concerned with advancing a general theory of law. He elucidated, in his Postscript to "The Concept of Law", that:

> It is *general* in the sense that it is not tied to any particular legal system or legal culture, but seeks to give an explanatory and clarifying account of law as a complex social and political institution with a rule-governed (and in that sense "normative") aspect. This institution, in spite of many variations in different cultures and in different times, has taken the same general form and structure, though many misunderstandings and obscuring myths, calling for clarification, have clustered round it (1994, pp. 239-240).

In this dissertation, so far, I have looked almost exclusively at the intuitions of lawyers and laypeople from one country: Brazil. Even if we count Turri & Blouw's data, all of our experiments come from only two countries. If my goal is to offer insight into a *general* theory of law in the Hartian sense, this won't do. I need to show whether there is significant cultural variation in the concept of rule. If there is, is it even possible to formulate an interesting general theory about rules? If there isn't, what explains the convergence? Moreover, how do these two questions interact with the different recognitional communities? These are the issues to which I turn in the next chapter.

4.

Rules around the world

Brian Tamanaha criticized Hart's analysis of the concept of law for being too parochial. According to the author, undue focus on a select set of legal systems have led Hart and other theorists before him to end up with an excessively narrow concept (Tamanaha, 2001). Last chapter's conclusions are open to a critique along these lines: the evidence surveyed might be important to a *particular* jurisprudence concerned with Brazilian law, but it is way too local to significantly advance debates about general jurisprudence.

This is especially the case when we take into account the even more radical possibility that law might be too diverse a phenomenon to lend itself to the sort of philosophical analysis that general jurisprudents do. This is a possibility entertained by Schauer (2015), for whom the existence of widespread cultural variation in people's attitudes towards central aspects of rules, such as the need for coercion, "may doom the very process of trying to find very much about law itself that is not culture-specific". He goes on to say that "There is no reason that law must have a cross-cultural essence and 'law' may merely be the label attached to a diverse collection of sociogovernmental phenomena neither joined by shared properties nor interestingly connected across different systems" (p. 74) In other words, it might be a mistake to analyze "the" concept of law. Perhaps, it would be much more interesting to drop the universalizing ambitions of analytic jurisprudence and focus on more informative analyses of particular legal systems, or, at least, clusters of similar systems.

This poses a challenge not only for the relevance of Studies 1-7, but for general jurisprudence as a whole. General jurisprudence presupposes that people everywhere share one single thing called law. Granted, law in Saudi Arabia might be very different from law in the U.K., but there is still something that makes both of them law, and that distinguishes both from other things, such as morality. This need not entail any old-fashioned analytical essentialism. Consider an analogy with the concept of game. The fact that no one could ever come up with a list of necessary and sufficient conditions that are shared by all games and only by games doesn't mean that there couldn't be a general theory of games that is interesting and informative. This general theory could expound on the typical, or most important, features of games, and draw the implications that these features have for those who want to engage with games (i.e., what general strategies can be employed to succeed in most games?), or simply understand its broader societal role (i.e., what features are associated with societies that play more or less games?). On the other hand, a general theory of games would be impossible – or, at the very least, very uninteresting – if there were no central features that most people everywhere would recognize as important for games. The fact that we can say interesting things about all games is evidence that even though this concept lacks an essence in terms of necessary and sufficient features, there is still some general concept of game that is distinctive, interesting, and informative. Something similar might be said of law[83]. If there are interesting properties that are particularly important (even if not necessary) for law everywhere, there is something interesting and informative to be said in general jurisprudence. In the event that there isn't, however, philosophers should refrain from these overarching claims and focus on more fruitful analysis of particular legal systems or clusters of legal systems.

[83] To see how this kind of theory is possible, but not interesting, consider Waldron's (2009) analogy with the concept of football. The word "football" is used to refer to one sport in the U.S. (the one played in the NFL), another one in Europe, Africa and South America (known by Americans as soccer), and yet another one in New Zealand (where, according to Waldron, it was used to refer to rugby). These three sports are radically different, but we can still imagine a general theory of football that could establish very abstract and mostly uninteresting propositions like "All football involves rules governing the enterprise of trying to move a ball from one end of a field to the other, against the opposition of a team trying to move it in the other direction" (p. 677). Surely, legal philosophy might have something slightly more interesting to say – e.g., about the structure of law as the union of primary and secondary rules – even if people's attitudes towards rule vary widely from culture to culture. Nonetheless, the point stands: general jurisprudence is *more* interesting if there is cross-cultural convergence around people's attitudes to rules.

So, empirical evidence about one single culture (or even a small sample of cultures) is not enough to establish that interesting philosophical theses about law, such as TP and WP are either true or false. It is possible that neither is universally – or even generally – true. In that case, we may fare better asserting for each legal system whether TP or WP holds than by saying less interesting – but more abstract and perhaps universal– things like CSM. Ultimately, whether or not a general jurisprudence of claims like TP and WP is possible or not is an empirical question.

But even those who are much less skeptical of general theories might object to the philosophical relevance of the evidence reviewed in chapter 2. After all, TP and WP are claims about the interpretation of rules, full stop. They are not claims about the interpretation of rules in Brazil and the U.S. Evidence from those countries counts in a broader effort to establish that those theses are either true or false, but they can only tell a very partial and incipient story.

I wholeheartedly agree with such criticism. Making experimental *general* jurisprudence necessitates cross-cultural investigation. There is a lot still to be done in this regard, but we have already begun the effort to collect systematic data about law in various cultures. As part of an experimental jurisprudence cross-cultural study swap led by Ivar Hannikainen[84], we joined forces with researchers from ten different countries to replicate Study 4 and investigate how well our conclusions generalize outside Brazil (Almeida et al., in prep).

First, we gathered cross-cultural data about the intuitions of laypeople around the world (Study 8). As I will report in the following section, Study 8 revealed significant cultural variation in the lay concept of rule: in some cultures, rules are interpreted in an almost exclusively textual way, while in other cultures they are seen as predominantly purposive. However, as Study 7 revealed, lawyers and laypeople diverge sometimes. Not only that, but that divergence is of jurisprudential significance. Hence, an appropriate cross-cultural evaluation of the concept of rule in the context of general jurisprudence must also include an appraisal of variation between lay and professional populations. To provide such appraisal, we selected a sample of countries lying at different points in the spectrum

[84] Available at: https://osf.io/sk7r3/

that goes from most textualist to most purposivist and surveyed the intuitions of lawyers through the same experimental design (Study 9). The results show that the professional concept of rule, unlike its lay counterpart, does not vary widely with culture. Finally, I consider the implications of the results for the viability of general jurisprudence and, specifically, for the theses of the Hart-Fuller debate.

4.1
Study 8 - Laypeople around the world (Almeida et al., in prep.)

Motivated by the interesting results and implications of the research reviewed in chapter 3, we set out to replicate it internationally. We used a simplified version of Study 4, which contrasted over- and underinclusion cases. Instead of combining this manipulation with all four conditions of the original experiment ("no prompts", "semantic prompt", "moral prompt", and "both prompts"), we dropped the conditions that introduced each part-question separately ("semantic prompt" and "moral prompt"). As a result, our cross-cultural study includes only the two conditions ("no prompts" and "both prompts") that differed the most in Study 4. As I have reported above, on that sample composed mostly of legally trained Brazilians, participants in the "both prompts" condition were much more willing to say that the rule was broken in cases of overinclusion than in cases of underinclusion. In contrast, they were equally likely to rate each kind of case to be a rule violation in the "no prompts" condition.

In light of our earlier results, we hypothesized that people around the world would show a similar pattern of results. To be precise, we believed that people would lean textualist (as predicted by TP) and that they would show the same asymmetry between judgments under the "no prompts" and "both prompts" condition detected in Study 4.

Thanks to the collaboration of several scholars[85], we collected responses in 11 countries: Brazil (207 participants)[86], Colombia (259)[87], Finland (142)[88],

[85] All collaborators will be invited to co-author the final draft. This invitation, however, is still in the future tense, so the list of authors will likely change until publication, both because some authors might decline the invitation, and because some new venues might be added, thus increasing the list.
[86] Provided by Noel Struchiner, Ivar Hannikainen, and myself.
[87] Provided by Alejandro Rosas Lopez.
[88] Provided by Michael Laakasuo.

Germany (359)[89], India (254)[90], Italy (350)[91], Lithuania (191)[92], Netherlands (391)[93], Poland (271)[94], Spain (286)[95], and the U.S. (254)[96]. Stimuli were translated from English to each local language by the scholars responsible for data collection.

Much to our surprise, we have found significant cross-cultural variation in judgments about rule violation in over- and underinclusion cases. Participants from some countries (Poland, Lithuania, Finland, Germany, and Italy) behaved in a textualist manner, judging overinclusion cases to be in violation of the rule more often than underinclusion cases. However, participants from a number of countries (Brazil, U.S., India, and Colombia), didn't show any detectable preference. Moreover, laypeople from the Netherlands and Spain actually behaved in a decidedly purposivist way, judging underinclusion cases to violate the rule significantly more often than overinclusion cases[97]. Overall, we failed to find the predicted asymmetry between the "no prompts" and "both prompts" conditions. Laypeople in most countries didn't seem to be affected by the prompts in the same way that our Brazilian sample did.

If we assume that the relevant recognitional community includes laypeople, these results are relevant for general jurisprudence. As foreshadowed by Schauer and Tamanaha, there is, indeed, substantial variation in people's attitudes towards law around the globe. Hence, most of the theses we surveyed in the preceding chapters hold only in some places. For instance, TP clearly holds in Lithuania, Finland, Poland, and Italy, while WP is true in all other countries. Under this broad conception of the recognitional community, this finding sheds doubt on the prospects of general jurisprudence. While it is true that some very general and abstract things can be said about rules in all of those countries (e.g., CSM), these

[89] Provided by Maruks Kneer.
[90] Provided by Vilius Dranseika, Markus Kneer, and Ivar Hannikainen.
[91] Provided by Alice Liefgreen.
[92] Provided by Vilius Dranseika.
[93] Provided by Niek Strohmaier.
[94] Provided by Piotr Bystranowski, Bartosz Janik, and Maciej Próchnicki.
[95] Provided by Ivar Hannikainen and Fernando Aguiar.
[96] Provided by Kevin Tobia.
[97] These lists were created after performing a student's t-test for the difference between the means in over- and underinclusion cases for each country. If there was a significant effect in either direction, countries were grouped as mostly textualist or mostly purposivist. Countries for which the difference fell above the 95% significance level ($p > .05$) were classified as not showing any detectable difference. This clarification is needed because the trend towards textualism noted among Brazilian participants was just slightly above the threshold ($p = .051$).

things are not particularly interesting or illuminating. If we focus on each country, however, and turn towards some sort of *particular* jurisprudence, we can start to gain some interesting insight into the way people engage with rules.

All of those conclusions are predicated upon the relevant recognitional community being comprised of laypeople. If we answer the recognitional question differently, by proposing that either judges or officials form the relevant community, then the results are less impactful, for we were simply looking for answers in the wrong place. This offers one good reason to run the same cross-cultural experiment again, but this time recruiting only lawyers.

There is also an additional reason to survey lawyers. Maybe, the divergence between Study 4 and this first cross-cultural result is not only driven by the quirks of Brazilian culture, but also by the populational differences between participants in each study. Study 4's sample was dominated by lawyers (they were 235 out of the 364 participants, or 65.56% of the total), while the vast majority of participants recruited outside Brazil were laypeople. Just as lawyers in Brazil diverge from laypeople about rule identity judgments (Study 7), these populations might also diverge elsewhere about rule violation judgments.

4.2
Study 9 - Lawyers around the world (Almeida et al., in prep.)

For our follow-up study, we hypothesized that even though the concept of rule varies widely among laypeople from different countries, it would be one single – and mostly textual – concept for lawyers. Not only that, but we posited that there would be a significant difference between the "no prompts" and "both prompts" conditions among lawyers. To test these ideas, we recruited lawyers from countries where laypeople showed a strong preference for textualism (Poland[98] – 161 lawyers and Finland[99] – 124), no discernible preference (the U.S. [100] – 159), and a significant preference for purposivism (the Netherlands[101] – 331). The study design and stimuli were exactly the same as in Study 8. Our hypotheses predicted that cultural

[98] Provided by Piotr Bystranowski, Bartosz Janik, and Maciej Próchnicki.
[99] Provided by Michael Laakasuo.
[100] Provided by Samantha Bensinger.
[101] Provided by Niek Strohmaier.

differences among professional populations should be much smaller than those between laypeople. Moreover, we should be able to replicate the findings of Study 4 about the effects of the prompts for textual and moral appraisal.

The results confirmed our hypotheses. Although some cultural difference remains, participants in all countries judged overinclusion cases to violate the rule significantly more often than underinclusion cases. Additionally, just as in Study 4, lawyers in this cross-cultural study were more aggressively textualist when both prompts were present than when no prompts were presented[102].

The implications for general jurisprudence are important. Convergence among lawyers, coupled with a restricted recognitional community, is sufficient to offer an empirical answer to Schauer's skeptical challenge against general jurisprudence. If the beliefs and attitudes of officials (or judges) are the crux of the matter, it is enough that *they* share one single concept of rule in order to make general jurisprudence something viable and interesting. Moreover, TP seems to hold steady among lawyers even in countries such as the Netherlands and the U.S. where laypeople behave as predicted by WP.

Finally, Study 4's counterintuitive and interesting results regarding the effects of analytic vs. holistic thinking successfully replicated among lawyers, but not with laypeople. One possible explanation is that in order for there to be an effect of analytic vs. holistic thinking, people must have previous conscious knowledge of the components being analyzed. In other words, if all their lives laypeople employed rules in a holistic fashion, the mere mention of text and purpose is insufficient to nudge them towards any direction. However, for those – like lawyers,

[102] There is one important methodological caveat. Our categorization of cases as over- or underinclusion depends on assessments of whether text and purpose were violated in each situation. By definition, an underinclusion case is one where people think that the text was not violated, but purpose was. However, participants presented with the underinclusion case in one of the scenarios we used, involving a rule prohibiting the use of phones inside a classroom, disagreed with our classification as measured by the supplementary probes introduced in the "both prompts" conditions. In some countries (Netherlands and Poland), people evaluated this vignette to be a core case, violating both text and purpose, while in others (the U.S. and Finland), people thought it was a borderline case with regards to text, even though it clearly violated the purpose. We interpreted this results as a failed manipulation check. Thus, we excluded this specific scenario from the analysis. If we include it back in the analysis, there is still convergence with regards to textualism among lawyers, but no statistically significant effect of "no prompts" vs. "both prompts".

but maybe not only lawyers – accustomed to dealing with rules in a more careful way, making text and purpose salient is enough to nudge them into textualism.

It is interesting to note that this result is inconsistent with ISA. If "Joan broke the rule" entails an indirect speech act to the effect that "Joan is blameworthy", this is a feature of natural language that should affect laypeople. The fact that it doesn't (Study 8) suggests that ISA cannot provide a complete explanation of Study 4's results.

4.3 Limitations

Even though the extension of our findings to almost a dozen countries goes a long way towards making them more general, there is still a sense in which our results are parochial. After all, apart from India, we only collected data in Western countries with very developed legal systems. Will the findings replicate in countries with different sorts of institutions? This is an important question for those, such as Tamanaha (2001), who want to extend jurisprudence beyond the Western cases that the literature usually takes as paradigm. My hypothesis is that we would likely see laypeople in such societies leaning towards the purposivist end of the spectrum, but if there are officials tasked with making and applying rules, I would expect them to behave in a more textualist manner. This, however, is nothing but an untested hypothesis based on my own possibly idiosyncratic intuitions. New studies focusing on people living under legal systems that depart from the Western paradigm are needed to test whether Tamanaha's intuitions are vindicated.

Another important limitation of this cross-cultural effort is that the sampling method varied in important respects from country to country. In Brazil and Poland, for instance, we collected data through snowball sampling, by posting the survey link on social media and asking colleagues, friends, and students to participate and spread the link. Given that all authors are affiliated with prestigious universities within their countries, we can expect that this sampling method will yield a very skewed sample of the population. The socioeconomic and educational status of respondents is probably much higher than that of the general population. Thus, it would be a stretch to think that these results are representative of how all laypeople

in each country think about rules. In stark contrast to this method, employed to gather data in countries including Brazil and Poland, in Spain participants were recruited through a specialized service (netquest) that guaranteed a sample that was representative of the general population. Future research should employ these more rigorous methods in all countries.

Nonetheless, we were still able to detect significant variation among countries with similar sampling methods. For example, we have reason to think that the samples collected in Poland, Brazil, and Colombia were all skewed in the same direction, but these three countries ended up in very different places in the purposivist-textualist spectrum. Likewise, samples in Germany, Spain, and the U.S. were provided by paid services that are not systematically skewed in any known direction; nonetheless, there was significant variation between the samples of those countries. Hence, although the diversity of sampling methods makes it harder to compare the results from country to country, we still found evidence of cultural variation, albeit in these restricted comparisons within groups of countries where similar sampling methods were employed.

4.4
Cross-cultural lessons about the concept of rule

In this chapter, I reviewed the findings of Almeida et al. (in prep). First, I argued that the very nature of *general* jurisprudence recommends a cross-cultural approach. Hart and Fuller were not primarily concerned with descriptions of particular legal systems. Their theses (in the context of this dissertation, CSM, TP, SP, and WP) were supposed to hold in any human society dealing with rules. But is it possible to make interesting claims that hold across all legal systems? Tamanaha and Schauer are skeptical. They advise that theorists should be open to the possibility that the diversity of law is such that the quest for one single concept of rule might be misguided.

In order to answer the skeptic challenge, experimental general jurisprudence must show that at least one recognitional community has robust intuitions across cultures. Study 8 showed that laypeople do not share the same set of intuitions in all countries, but Study 9 revealed that lawyers do. So, if we are aiming at a general

theory of rules (and, inasmuch as law is centrally about rules, a general theory of law), we should focus on some more restricted recognitional community, and not on laypeople.

As to the theses, TP and WP hold among different populations. Laypeople in some countries are mostly purposivist, while in others they are mostly textualist. On the other hand, lawyers are always textualists. Finally, the fact that the contrast between the “no prompts” and “both prompts” conditions obtains only among professionals provides an additional data point to help understand the nature of the effects.

5.

New questions

Throughout the last three chapters, I have reviewed the evidence regarding rule-related judgments in the context of questions that are pervasive in the philosophical literature about law. Studies 1-9 provide data about claims that are central to those debates. In that sense, the preceding chapters try to apply new tools to old problems. But one of the interesting things about pursuing philosophical ideas through experiments is that the process of designing studies and analyzing the results can also suggest new perspectives. Thinking about which scenarios are able to test a specific philosophical claim sometimes reveals that some overlooked aspect of the investigated concept is in need of more attention. In a similar way, thinking about the implications of unexpected results often suggests new hypotheses that might provide some insight into the concept we're studying.

In this chapter, I will pursue two sets of questions about rules that become salient once we adopt the standpoint of experimental philosophy. Even if they are not entirely new, they have received significantly less attention than other aspects of the philosophical debate about the nature of rules.

The first set of questions pertains to purposes. The Hart-Fuller debate is full of references to purposes, but contains no explicit attempt at a definition of such concept. Hart speaks vaguely of "social aims", while Fuller obliquely alludes to the fact that purposes are moral in nature. However, in order to test specific hypotheses about purposes (e.g., is WP true?) one needs to spell out these views in more detail. Doing so reveals interesting questions and ambiguities. For instance: what is the

relationship between purposes and the intentions of rule-makers? Is current practice relevant in determining purpose? What happens when the intended purpose is evil? Drawing from the psychological literature and from the findings of Almeida, Knobe, Hannikainen & Struchiner (in prep.), I will explore those questions and point towards interesting paths for future research.

The second theme relates to whether the explanation for Studies 1-9 lies in psychological states – transient attitudes that might change from one moment to the other within one single person – or in personal traits – that vary cross-sectionally from person to person and are more persistent through time. Entertaining the hypothesis that personal traits might help explain Studies 1-9 reveals four distinct decision-making profiles defined as attitudes to Schauer's categories of cases. People might be hardcore textualists, rule maximalists, rule minimalists and passionate purposivists. Different normative virtues and vices underlie each profile, but the literature about legal decision-making overlooked some of them (rule maximalists and rule minimalists) so far.

5.1
Purpose

In his review of "The Morality of Law", Hart commented that:

> the virtues and vices of this book seem to me to spring from the same single source. The author has all his life been in love with the notion of purpose and this passion, like any other, can both inspire and blind a man. I have tried to show that I would not wish him to terminate his longstanding union with this *idée maîtresse*. But I wish that the high romance would settle down to some cooler form of regard. When this happens, the author's many readers will feel the drop in temperature; but they will be amply compensated by an increase in light (Hart, 1983, p. 363).

In private correspondence, Fuller took up the romantic imagery and conceded to some obscurity:

> All I can say of Miss Purpose is that the Old Girl still looks good to me. One of her enduring charms is that she is a very complex creature indeed, subject to unpredictable moods of surrender and withdrawal. I believe deeply in her without pretending that I really understand her (quoted in Lacey, 2010, p. 40).

While Fuller's lack of precision certainly provided ammunition for Hart in their philosophical debate, a lack of precision in experimental affairs can be much

more costly. In Studies 1-9, morally suspect results were always also putatively violations of the original intentions of the rule-maker. For instance, under the "no shoes" rule, presumably both the normative goal of "keeping the apartment clean" and the intentions of the rule-maker narrowly construed effectively allowed people to walk in with pristine shoes. This goes to show that the original intentions of the rule-maker and a moral view of purposes may sometimes coincide. But those two factors might also come apart. Consider the following scenario:

> **Housing regulations**
>
> The city council of Santa Cruz, a city in Brazil, decided by unanimous vote to institute strict housing regulations. A number of slums were forming in the city's hillsides and the bill mandated the removal of all hillside houses and the prohibition of further development on hills. Although every single one of the 50 city counselors supported the bill, different groups of lawmakers voted for this law for very different reasons. 30 lawmakers voted for the bill because they wanted to increase the value of the land near Santa Cruz's downtown, which was depreciating as a result of the slums. The remaining 20 counselors supported the bill because they wanted to protect the poor people living in the slums from landslides, which is something that often takes place in Santa Cruz's hills.
>
> The outcomes of the law show that both objectives were at least partially accomplished. Reports of deaths and injuries resulting from landslides fell significantly, and the value of land near Santa Cruz's downtown has increased back to pre-slum heights.
>
> Now that the law is in effect, courts and legal scholars are inquiring about what is the purpose underlying it. After all, the purpose of a rule is often used in a legal setting to decide whether someone is legally liable (Almeida, Knobe, Hannikainen & Struchiner, in prep.).

In this scenario, the original intentions behind the regulation are morally suspect, but the same bill also achieves morally laudable ends. Which of those two different normative goals is the purpose of the rule? The morally good goal of protecting slum dwellers, or the morally worse goal of increasing the value of land downtown? Legal theorists disagree as to the correct answer. Diverging answers, by their turn, might carry important implications for the theoretical conclusions regarding SP and WP.

Fuller's theory seems to suggest that, in this case, the moral goal pursued by the minority is the purpose of the legal rule. Recall the tension between SP and the Separability Thesis[103]: Fuller's brand of purposivism represents a challenge to

[103] Section 2.2.

positivism precisely because it argues that interpretation is necessarily *moral*. If, in contrast, purposes could be determined on purely factual basis, there would be no incompatibility between purposivism and positivism. Hence, his empirical prediction would presumably be that people's interpretations of the rule would be guided by the goal of protecting people from landslides, and not the goal of increasing the value of land. To ascertain the importance of purpose, in that context, is to argue that hillside houses under no risk from landslides (overinclusion cases) do not violate the rule, and that houses that are not in hillsides, but nonetheless suffers from the possibility of landslides (underinclusion cases) do infringe upon the law.

This is clearest from Fuller's rejection of the alternative view that the intentions of rule-makers are all that matter. His resistance to this idea is perhaps most transparent in his discussion of the disanalogy between law and military orders. For Fuller, those two domains demanded very different interpretative strategies. If legal rules were anything like military orders, then "the task of the interpreter [would] be to discern as best he can the desires of the high command" (1969, p. 229).

> But deciding what the legislature would have said if it had been able to express its intention more precisely, or if it had not overlooked the interaction of its statute with other laws already on the books, or if it had realized that the supreme court was about to reverse a relevant precedent – these and other like questions can remind us that there is something more to the task of interpreting statutes than simply "carrying out the intention of the legislature" (p. 231).

In contrast, to interpret the law properly, people must "put themselves in the position in which the accused found himself and ask what can *reasonably* be expected of a human being so placed" (p. 229, emphasis added). We called this position *moral purposivism* (Almeida, Knobe, Hannikainen & Struchiner, in prep.). According to it, the purpose of a rule is the morally tinted goal it pursues. Moral purposivism is also transparent in Hart & Sacks' famous remark that legal interpreters should assume "that the legislature was made up of reasonable persons pursuing reasonable purposes reasonably" (Hart & Sacks, 1994, p. 1125). Shecaira (2009), drawing from Dworkin, offers a more precise formulation of this idea, stating that moral purposivists attribute to the rule "the purpose that best justifies it from the standpoint of political morality" (p. 46).

A different set of legal theorists completely disagrees with Fuller's disanalogy. For this strand of philosophers – for instance Alexander & Prakash (2004), Alexander & Sherwin (2008), Knapp & Michael (1982; 2005), and Fish (2005) – the task of interpretation is to devise the intentions of rule-makers[104]. They go so far as to state that "[intention free] textualism is a conceptual impossibility" (Alexander & Prakash, 2004, p. 969)[105]. All meaning – within and without the legal domain – is a function of the intentions of the speaker, understood as a historically situated set of mental states regarding their speech acts. Alexander & Sherwin (2008) propose the following example as an appeal to our intuitions:

> […S]uppose [your] Mom has never mastered the distinction between autobahn and ottoman, and she leaves you a note requesting that you pull up the "autobahn" next to the sofa when she comes to visit. You surely know what to do, and it isn't to run a German highway through your den (p. 133).

In order to correctly interpret the note, the son must understand what his mother had in mind in writing the note. But what about purposes? Shecaira (2009) and Barak (2005) suggest that such focus on intentions can be, and often is, applied to the search of purposes. Just as intentionalist interpretation of the note requires attention to the mental states that animated note-drafting, purposive interpretation of a rule under that outlook requires us to consider not the purpose that best justifies the rule from a moral standpoint, but the one that rule makers originally had in mind, regardless of its moral standing. Thus, according to this view – which we dubbed *"original intentions" purposivism* – the purpose of the "housing regulations" rule is the one intended by the majority. Note that – unlike moral purposivism – even very strong versions of "original intentions" purposivism are

[104] The intentionalists' claims aim more squarely on whether rule makers intended specific situations to fall within or without the scope of legal rules (see Shecaira, 2009; Alexander & Prakash, 2004, Appendix II). This narrow version of intentionalism that is centered on what gives meaning to legal rule's texts actually competes with CSM, and not with the purposivist thesis. According to it, CSM is false because interpretation requires necessary reference to the intentions of the speaker, so text is never sufficient. Aware of this narrower focus, intentionalists such as Alexander & Sherwin (2008) distinguish between intentions and purposes (see p. 142; Alexander & Prakash, 2004, Appendix II). Hence, the sort of purposivism that focus on the original intentions of rule makers, while certainly inspired by intentionalist ideas, is not necessarily a stance that this set of authors must subscribe to. Instead, the argument is just that the same methods that intentionalists use to ascertain meaning might be leveraged at a higher level of abstraction to uncover the originally intended purposes.

[105] For a convincing critique of this strong thesis, see Sinnott-Armstrong (2005).

still compatible with the Separability Thesis, for we might be able to discover the mental states held by rule-makers without having to resort to any moral criteria.

This shows how WP is ambiguous in a very consequential way. If an intentionalist reading of WP is available, the evidence gathered in Studies 1-9 is ultimately unable to deliver Fuller's natural law outlook any definite win: in every case where WP fares better than TP, it might be because of the original intentions of rule-makers, and not because of the moral force of purposes[106].

The only way to remedy these confusions is to gather more evidence. For instance, what does existing experimental work about purpose attribution have to say about these two competing views? If most people within a recognitional community equate purposes with original intentions, then the results bearing to WP should be read in an intentionalist way. Alternatively, if most people are moral purposivists, the partial success of WP should entail Fullerian conclusions. Finally, it might turn out that some other way of determining purpose might be relevant. In that case, we might have to consider yet another version of purposivism and its corresponding consequences to the four theses under discussion.

Thankfully, purpose attribution is not something that interests only lawyers. Cognitive scientists have long studied how people attribute purposes to physical objects. As analogies such as the one Fuller draws between rules and inventions suggest, there might be an important connection between purposes in law and in life. Perhaps, surveying existing research on cognitive science might provide insight into legal purposes.

5.1.1
The psychology of purpose attribution

The most prominent strand in the psychology of teleology assumes that purposes are a matter of the intentions of a designer. According to this view, the prototypical thing that has a telos is an artifact: artifacts are objects built by humans

[106] Which doesn't mean that Studies 1-9 tell us nothing. They clearly establish that text prevails over *both* moral purposes and original intentions in some circumstances and within some populations. Moreover, the finding that either the normative goals originally pursued by rule makers or that moral purposes play into the very concept of rule is already sufficient to establish an important connection between law and morality, even though it may be contingent and circumscribed to some recognitional communities.

to achieve goals; hence, the purpose of an artifact is just whatever goal its designer was trying to achieve. Thus, psychological orthodoxy in teleology shares many features with "original intentions" purposivism. Sometimes, this view makes intuitive sense, as in the case of artifacts, and, perhaps, of biological structures, such as noses and eyes (which were presumably adaptive because of the goals they achieved). However, it does not extend easily to other kinds of things, such as mountains and clouds. Surely, if purposes are all about the intentions of a designer, those things can't have purposes.

Hence, Kelemen's (1999) results struck the field as very surprising: she found that preschool children employed a "promiscuous teleology", attributing purposes to inanimate natural objects that lacked a designer (e.g., saying that "clouds are for raining"). In later work, Kelemen (2004) argued that this provided evidence that children are "intuitive theists"; they intuitively endorse the existence of some non-human designer who created these inanimate natural objects. Thus, their teleological talk about these inanimate natural objects makes reference to the intentions of those supernatural beings. Later work showed that adults without exposure to formal science education showed similarly promiscuous teleological judgments (Casler & Kelemen, 2008), as did adults pressed for time (Roberts, Wastell & Polito, 2020).

A similar understanding of purpose animates the work of David Rose. In a series of recent papers, Rose and colleagues showed that reasoning about the composition (Rose & Schaffer, 2017) and persistence (Rose, 2015; Rose, Schaffer & Tobia, 2018) of objects, as well as the essence (Rose & Nichols, 2019; 2020) of artifacts and animals are driven by teleology. He endorses the "promiscuous teleology" view, arguing that our intuitions about mereology are misguided precisely because they presuppose non-existent design.

This goes to show that there is plenty of support for the idea that people equate purposes with the intentions of a designer in psychological research. When asked what a thing is for, children often report the intentions of the thing's designer. Adults keep on doing this throughout, but become less and less promiscuous, restricting teleological explanations to artifacts and biological entities, at least when scientifically educated.

In contrast, no work in cognitive science explored the hypothesis raised by moral purposivists. So, even though people often identify the purposes of physical objects with their originally intended goals, there is no evidence regarding how they would behave in cases such as "housing regulations". Moreover, there is an important gulf between extant psychological work on teleology and the questions about rules that I want to answer. First, all of the evidence in the studies cited in the last two paragraph deals with tangible objects. Rules are intangible. While we have reason to think of rules and artifacts as similar in important senses (e.g., both are created by someone who wants to achieve some goal that might be morally good or bad), they are also dissimilar in other relevant ways (e.g., artifacts often achieve their function by possessing some physical properties, while rules do so by other means[107]). Second, this literature focuses on teleological *explanations*, while we are interested in other judgments about rules. Specifically, we are interested in questions about the rule's identity and about rule violation.

5.1.2
Study 10 - Original intentions and moral purposes (Almeida, Knobe, Hannikainen & Struchiner, in prep.)

To address these issues, we (Almeida, Knobe, Hannikainen & Struchiner, in prep) ran a series of experiments. One of them built upon the structure of the "housing regulations" scenario: participants were randomly assigned to a condition where either a majority or a minority of lawmakers supported the morally good purpose. After reading the scenario, participants then answered whether they agreed with a statement about the purpose of the rule. Half of the participants received a statement that the morally bad purpose was the purpose of the rule, while the other half evaluated a sentence affirming that the morally good purpose was the purpose of the rule.

The results once more revealed a discrepancy between lawyers and laypeople. Participants without legal training, recruited in the U.S., showed a strong

[107] Some of the literature see this as a relevant difference. Kelemen (2006) notes that earlier work failed to show an effect because "experimenters unnaturally dissociated artifact form from artifact function", and that the approach of "using artifacts that look designed in that their structural properties clearly relate to their functional affordances" is a better test for the promiscuous teleology hypothesis.

preference for the purpose intended by a majority of lawmakers, and no preference at all for morally good purposes. Brazilian lawyers, in contrast, were affected by both factors. They agreed more often with the statement attributing a purpose to a rule both when the purpose was supported by the majority and when the purpose was morally good.

Hence, Study 10 provides further evidence that original intentions shape purpose attribution. Both lawyers and laypeople were more willing to attribute purposes supported by a majority than those supported by a minority. But we also found evidence that purpose attribution by lawyers in the legal domain is shaped by morality – just as Fuller and company would have predicted[108].

How does that result interact with the evidence surveyed in the preceding chapters? The fact that lawyers and laypeople diverged suggests a separate set of implications for each group. Here, the suggestion is that Fuller's version of the purposivist theses surveyed in chapter 2 might perform better as a description of the way lawyers devise legal purposes. On the other hand, a purely intentionalist view might provide the best explanation for the way laypeople go about attributing purposes. This divide is consequential: purposes more often than not are left implicit. If lawyers let their search for a rule's purpose be informed by political morality in a way that laypeople do not, both groups are bound to disagree about purposive interpretation. Moreover, here Fuller lacks a way to save face: if he really had laypeople as his target recognitional community, as Rundle (2012) argues, then his descriptive account of how they behave seems to have been simply misguided.

The situation, however, is even more complex, as recent studies show that yet another factor shapes purpose attribution judgments: current practice.

[108] There are actually two potentially relevant differences between the two populations: one was recruited in Brazil, while the other was recruited in the U.S., and one is comprised of lawyers, while the other of laypeople. I am assuming that the relevant difference is that between lawyers and non-lawyers. The assumption is based on previous findings reported in chapter 4 that there is widespread convergence between lawyers in different jurisdictions. This assumption is, however, defeasible. Naturally, the best test would be to conduct a similar experiment comparing laypeople and lawyers from the same country.

5.1.3
Study 11 – Current use and moral purposes (Almeida, Knobe, Hannikainen & Struchiner, in prep.)

Joo, Yousif & Knobe (2020) showed that people often identify an object's purpose with the goal it currently serves. For instance, when asked what's the purpose of metal tubes that were originally intended as straws, but are now being used as windchimes, people agree with the option stating that "The metal tubes are for making music when the wind blows". In other words, under at least some circumstances, current practice trumps the original intentions of the object's designer as a way to determine purpose. Further experiments in the same paper suggest that this is due in equal amounts to the explicit recognition of members of the relevant community that the object's purpose has changed (i.e., people say things like "These metal tubes are now used as windchimes, and not as straws"), and to the social fact that behavior has changed (even without explicit recognition).

Recent work in legal scholarship also defends that current practice is very important in legal interpretation. Levin (2012) argues that the dominant approaches to legal interpretation, including textualism, moral purposivism, and intentionalism, "ignore or minimize the manner in which citizens actually determine what the law requires of them, namely, by observing the behaviors of other citizens and officials" (p. 1114)[109].

This raises the following questions: does the preference for current practice also apply to the attribution of purposes to legal rules? Or is it a quirk of how we think about physical objects? How do current practice relate to the influence of morality we observed among lawyers? Moreover, just as the ambiguity between moral purposes and original intentions proved consequential to the broader implications of SP and WP, the same is true about the ambiguity between moral purposes and current practice: data about actual practice is just as factual and non-

[109] Levin further emphasizes his point by stating that "This account seems obvious to me as a descriptive matter" (2012, p. 1116). Just as Alexander & Sherwin, Levin also seems preoccupied with a narrower understanding of interpretation that might be distinguishable from talk about purposes. Nonetheless, once again I believe we can treat it as a difference of degree. If current practice can be used to determine the narrow application of a statute to a case, it can surely also help us devise the broader, vaguer, and more normatively charged goals lying behind the rule's formulation.

moral as historical data about original intentions. Hence, a current practice reading of SP and WP need not entail any denial of the Separability Thesis.

To answer these questions, we (Almeida, Knobe, Hannikainen & Struchiner, in prep.) devised a set of vignettes where a rule changed purposes in a way that flipped its moral valence. In the "moral improvement" condition, a rule (e.g., stating that people must wear gloves) that was originally intended to advance a morally defective goal (e.g., to show support for a homophobic and misogynistic cult) is now used to achieve something laudable (e.g., stopping the spread of a lethal disease). In the "moral deterioration" condition, the only change is that the original intentions of the rule-maker were morally good, but became morally bad after some length of time.

In this design, participants showed a strong preference for current use in purpose attribution judgments. No effects of morality could be detected. Additionally, purpose attribution judgments were unaffected by whether the thing described was a physical object or a rule. The exact same pattern held for both lawyers and laypeople.

Summing up, there is evidence for at least three different kinds of purpose attribution to rules: (a) sometimes, people think that the rule's purpose is determined by the intentions of the rule-maker; (b) other times, people think that a rule's purpose is defined by current practice surrounding the rule; and, finally, (c) lawyers' purpose attribution is sometimes partly a function of the moral quality of the pursued goals. Each of these kinds of purpose entails different jurisprudential consequences. Only (c) poses a challenge for legal positivism. However, even the gulf between (a) and (b) as descriptive claims call for very different explanations. A positivistic legal theory focused on the original intentions of lawmakers is very different from one focused on current practice[110].

Under which circumstances do people prefer each of these three strategies for attributing purposes to rules? What contexts make people sensitive to one or the other kind of purpose? These questions are far from being definitely answered, but

[110] Contrast Alexander & Sherwin (2008) with Gardner (2012, chp. 4).

they should loom large as a research goal for jurisprudents concerned with the role of purpose in legal interpretation.

5.1.4
Further open questions about purposes

This is only one in a series of interesting questions about purposes that necessitate further empirical investigation. Consider the following issues and some tentative suggestions on how to test them.

In some cases, such as those investigated in Studies 1-9, the original intentions of the rule-maker coincide with moral purposes and with current practice. In those cases, we noted that purpose seemed to make a difference to rule violation judgments: people were overall more likely to say that a rule was violated when its purpose was also violated than when it wasn't. In other cases, however, those three factors diverge. Studies 10 and 11 help us know what factors influence people's views about purposes in those diverging cases. More specifically, they tell us that the three factors identified in the literature (original intentions, moral purposes, and current practice) exert influence under certain circumstances, and that current practice seems to prevail when in conflict with other factors. These results, however, concern judgments of purpose attribution, and not of rule violation. The fact that people identify the purpose of a rule with morally evil goals defined by intentions or practices doesn't tell us anything about whether they take these goals into account in decision-making. After all, even if they think that rules sometimes serve immoral purposes, do people let them influence their decision-making? When we are faced with heinous rules created and applied with evil intentions, without any chance of a redeeming interpretation, do we remain purposivists? Do we become textualists? Or do people simply let their direct moral evaluation of the situations at hand guide their judgment in such occasions?

Consider the following example: imagine a discriminatory rule barring persons with a given skin color from frequenting a given university. The rule's intended purpose is to restrict elite education to a privileged racial group. Current practice surrounding the rule reinforces the same normative goal. Clearly, this is an evil rule with an evil purpose. Moreover, it is very difficult (likely impossible) to

find a morally good justification for this evil rule. Would people still let this odious purpose inform their rule violation judgments?

To make things more concrete, imagine an underinclusion case under that rule: Martha doesn't have the prohibited skin color, but nonetheless descends from and belongs to the group discriminated against. She applies for the racist university. Does the rule recommend the denial of Martha's application? If non-moral purposes really influence rule violation judgments, Study 2 leads us to believe that a sizeable minority (~20% of respondents) would answer yes. Alternatively, if the right complement to the textual element of a rule is not purpose strictly speaking, but some more general moral judgment, almost no one should respond this way (treating it as they treated off-topic cases in Study 2).

A similar test involves using morally bad purposes to achieve morally good results. Imagine a discriminatory rule forbidding people from a certain caste from swimming in a public swimming pool. Suppose further that the originally intended and currently pursued purpose is to avoid interactions between members of different castes. Again, there is little doubt that the rule and its purpose are morally in the wrong. Similarly, it is hard to imagine a reading that gives this rule a good purpose. But imagine an overinclusion case under that rule: Peter, a member of the barred caste goes for a swim at 4 a.m., when absolutely no one else is in the pool. Is Peter in violation of the rule? Even though the purpose is evil, it may be invoked in that instance to achieve a morally good result: if no one else was around, this textual infraction certainly did not result in interaction between members of different castes. However, if morally bad purposes are only invoked to achieve morally good results, unmediated moral judgments, and not purposes, are doing all the work. In that case, a DNR account of rules should substitute the reference to "purpose" with one to moral judgments.

Another important question that is empirically testable regards implicit purposes. In Studies 1-11, one single purpose was explicitly stated. But this is seldom the case in real life. Ordinarily, we have access to a rule's text and the overall context of the situation, but not to detailed information about the mental states that went into the making of the rule and the normative goals that were pursued through it. This means that, in ordinary life, we often have to infer purpose

from incomplete information. But what purposes do people infer? Do they infer only one or multiple purposes?

The realization that rules might have multiple purposes is not new. Both Barak (2005) and Schauer (1991) acknowledge the possibility. More often than not, they are concerned with purposes in different levels of generality[111]. For instance, the "no dogs in the restaurant" rule has the immediate purpose of avoiding annoyances to patrons in the restaurant; but that immediate purpose is itself justified by deeper commitments, such as increasing restaurant profits, which by their turn might rest on even deeper normative goals, and so on (Schauer, 1991, p. 73).

While the existence of multiple purposes in that dimension raises a series of interesting issues, I am more interested in the dimension of multiplicity that Barak (2005, p. 113) calls horizontal: some rules might have multiple purposes formulated in the same level of abstraction. What are the consequences of these cases for legal theory? Despite both acknowledging this possibility, neither Schauer nor Barak consider it in detail. But this is an interesting question, because each purpose defines a different set of off-topic, core, over- and underinclusion cases.

Let's represent this by reference to a rule R that might be justified by either purpose A or purpose B. Imagine a vignette that is ambiguous as to whether the morally best/actually intended/currently pursued purpose is A or B, but that either purpose is plausible. Moreover, both purposes are different, but compatible: the requirements of purpose A, while different from the requirements of purpose B, do not conflict with them. Three distinct patterns might happen, as represented in Table 1:

	Purpose A	Purpose B
Pattern 1 Both purposes	Core = 100% Over = 60% Under = 20% Off = 0%	Core = 100% Over = 60% Under = 20% Off = 0%
Pattern 2 Only purpose A	Core = 100% Over = 60% Under = 20%	Core = 100% Over = 100% Under = 0%

[111] What Barak (2005, p. 113) calls vertical purposes.

	Off = 0%	Off = 0%
Pattern 3 Only purpose B	Core = 100% Over = 100% Under = 0% Off = 0%	Core = 100% Over = 60% Under = 20% Off = 0%

Table 1 - People might infer that a rule has either one of the purposes separately or both of them simultaneously. Estimates for the amount of people judging over- and underinclusion cases to violate the rule were approximated based on Study 2.

In other words, people might privilege one purpose over the other, or think that the rule has two distinct and equally valid purposes. In that case, we shouldn't speak of "the" purpose and "the" set of over- and underinclusive cases, but of several possible purposes and sets of over- and underinclusive cases[112]. Data from Almeida, Knobe, Hannikainen & Struchiner (in prep.) suggests that people are indeed willing to attribute multiple purposes to a single rule, even when questions are phrased in terms of "the" purpose. This prospect complicates the analysis of rules much further. If something like the multiple horizontal purposes view obtains, we might have to take several purposes into account before assessing whether the rule was violated or not.

Each of the puzzles I discussed above (are the purposes of a rule defined by original intentions, morality, or current practice? Are rule violation judgments influenced by morally evil purposes? What guides implicit purpose attribution judgments? Can rules have multiple purposes?) came up as a result of experimental investigations, either during the process of creating experimental vignettes, or as a reaction to surprising results. All of them are important for philosophical reflection about the concept of purpose, and, insofar as purposes are important for rule-related judgments, the concept of rule. Moreover, each entails testable hypotheses that might further our knowledge about these concepts.

[112] Schauer (1991) treats rule as a relational concept ("[…] the concept of rule may be seen to be more of a *relationship* than an isolated entity", p. 73): a rule is defined by the capacity of its instantiation (often, but not necessarily, its text) to resist its purposes. The possibility of multiple purposes at the same level of abstraction raises an interesting question about rule individuation in the Schauerian framework: if something like Pattern 1 obtains, are we confronted with one single rule (with a composite purpose that is the disjunction of purposes A and B) or with two distinct rules? Schauer isn't clear on that front.

5.2
Traits or states?

The data from Studies 1-11 might mean two different things. Maybe (1) some people are purposivists, while others are textualists, or (2) sometimes we feel the purposivist pull, while other times we adopt a textualist standpoint. We might say that the first interpretation treats purposivism and textualism as *traits*, while the second treats them as *states*. Traits are longstanding characteristics (either acquired or innate) that differ cross-sectionally *between* persons[113], while states are transient, changing from one moment to the next *within* each person. Which of the two accounts should we prefer as an explanation for the data? What implications are associated with each explanation?

While the successful experimental manipulations across the holistic/analytic divide in Studies 4 and 9 suggest we are dealing with psychological states (after all, if the questions we ask are enough to shape judgments, it is hard to think that they were based on very stable dispositions to begin with), the existence of populational differences in Studies 7, 9, and 10 pulls towards the trait-based explanation. Moreover, the trait hypothesis has the benefit of accounting for Hart and Fuller's own understanding of their disagreement: recall the quote from Hart reproduced in the introduction stating that maybe his intuitions were so different from Fuller's that he should never be able to understand him (1983, p. 343), and the remarks by Fuller quoted in chapter 3 rejecting the analytical style of thinking (Fuller, 1969, p. 191).

In any event, further evidence should try to adjudicate between those two alternatives by running longitudinal studies where the same group of people is asked to adjudicate a larger collection of core, off-topic, under-, and overinclusion cases at different points in time. At this moment, it pays to ask: What could the results of such an investigation reveal? If there are different decision-making profiles, which form do they assume?

[113] As Hudson & Fraley (2015) put it, "[…] a large body of research has demonstrated that, despite being relatively stable, personality traits are malleable and change in response to a variety of external factors, including normative life experiences" (sources omitted, p. 490). Recent research has shown that some relatively stable traits, such as people's empathy, are also subject to changes across time (Hannikainen et al., 2020).

A cursory glance over the Hart-Fuller debate and the literature on legal decision-making (such as Schauer, 1991) would lead people to postulate two distinct profiles, which I call hardcore textualists and passionate purposivists[114]. First, consider hardcore textualists. This group of people would take TP to the extreme, maintaining that text is always the only thing that matters as far as rule violation goes. How would they answer each category of case defined by the relationship between text and purpose? They would say that core and overinclusion cases break the rules, while underinclusion and off-topic cases do not. In other words, a hardcore textualist treats textual violation as a necessary and sufficient condition for rule application, while violation of a rule's purpose is neither necessary nor sufficient.

Now, picture the mirror image of hardcore textualists: passionate purposivists. They would say that core and underinclusion cases are violations of the rule, but not overinclusion and off-topic cases. For them, the roles played by text and purpose are exactly the opposite as those identified by hardcore textualists: violation of purpose is a necessary and sufficient condition for rule violation, while violation of the rule's text is neither.

Do hardcore textualists and passionate purposivists exhaust the decision-making space? Let's first consider this question as an empirical statement that can be tested against the backdrop of Study 2. On that experiment, every participant answered one case under each of Schauer's categories: a core case, an overinclusion case, an underinclusion case, and an off-topic case. If everyone participating in Study 2 was either a hardcore textualist or a passionate purposivist, the proportion of positive answers to over and underinclusion cases should sum to 100%. Imagine 60% of people are hardcore textualists. As hardcore textualists, they would say that overinclusion cases violate the rule and that underinclusion cases do not. If everyone is either a hardcore textualist or a passionate purposivist, the remaining 40% of respondents would have to be passionate purposivists giving flipped answers to these two questions.

[114] Schauer discusses formalists and pure particularists in similar, but not identical terms. A careful analysis of the similarities and differences between my typology and Schauer's is beyond the scope of this dissertation.

This, however, is not what we found. Only 93 of the 200 participants behaved as hardcore textualists (77) or passionate purposivists (16). The remaining 107 participants did not conform to these decision-making profiles. Why is it so? Are there other coherent decision-making profiles defined by the relationship between text and purpose? After discarding incoherent combinations that treat either core cases as complying with the rule or off-topic cases as violating the rule, we can see that there are two other coherent stances one can take on the relationship of text and purpose to rule violation judgments. Besides hardcore textualists and passionate purposivists, people might also be rule maximalists or rule minimalists, profiles that received much less attention from the philosophical literature.

Consider first the rule maximalist: for such a person, rules are a matter of text and purpose, but either one of them is sufficient to entail rule violation. The rule maximalist thus judges core, overinclusion and underinclusion cases all to be violations of the rule, reserving compliance only to off-topic situations. The other sensible position would be that of a rule minimalist. For them, rules are also a matter of text and purpose, but both must be simultaneously violated for the rule to be broken. Neither violations to text, nor to purpose, suffice to entail rule violation judgments. Instead, both kinds of violation are individually necessary and jointly sufficient to arrive at such a judgment. Thus, only core cases count as rule violations to rule minimalists.

These positions, as well as the number of participants in Study 2 that conformed to them, are presented in Table 2. It is important to note that Study 2's data is relevant only in a very restricted sense: if no one in Study 2 showed the pattern of results predicted by each profile, we would have good reason to suspect of their validity as an empirical explanation for the decisions people make. However, the fact that some people conformed to the predictions of such profiles is far from showing that the trait-based explanation is true. Study 2 was not designed to establish anything about whether those decision-making profiles are really picking out traits. The fact that the particular set of cases judged by a participant happened to conform to a profile might be perfectly well explained by some momentary disposition. Only longitudinal studies that focus on the same set of people over some amount of time might shed light on whether those profiles are in any way stable or not. Moreover, participants answered only one single case under

each category. Thus, even if these profiles really are picking out different outlooks towards rules that different people maintain, our data would likely be too noisy to provide much insight.

	Text	Purpose	Text & Purpose
Hardcore textualist (77)	Necessary Sufficient	Not necessary Not sufficient	--- ---
Rule maximalist (19)	Not necessary Sufficient	Not necessary Sufficient	--- ---
Rule minimalist (61)	Necessary Not sufficient	Necessary Not sufficient	--- Sufficient
Passionate purposivist (16)	Not necessary Not sufficient	Necessary Sufficient	--- ---

Table 2 - Coherent ways of treating text and purpose as conditions for rule application.

The existence of trait-like decision-making profiles is thus still only a hypothesis. But it is interesting nonetheless. Previous work had much to say about the relationships between text, purpose, and legal decision-making models. Yet, not all possible attitudes towards those relationships were explored with equal care. Thus, uncovering the hypotheses that people might be rule maximalists and rule minimalists might be a benefit afforded by experimental jurisprudence.

This hypothesis is especially interesting because the stances of rule maximalists and rule minimalists are normatively compelling. As such, it might be interesting to contemplate them, even if they turn out to be useless as empirical descriptions. Think of areas of law where judicial minimalism is celebrated as a virtue, such as criminal law, or tax law: one elegant way to describe what judges should do in such cases is to recommend that they behave as rule minimalists, reserving sanctions to core cases. This is true independently of whether judges working in this area are already rule minimalists. If they are not, we can argue that they should be; if they are, then we found an interesting and precise way to describe and justify their behavior. On the other hand, areas of law protecting the rights of underpowered parties, such as labor law or consumer law, should mandate that

judges behave as rule maximalists – again, regardless of their actual attitudes about text and purpose. Ample protection to these rights is so dear that either textual or purposive violation should be sufficient to justify punishment[115].

These normative arguments might actually even pull against the trait hypothesis: according to them, in countries such as Brazil, where judges are expected to hear cases on all areas of law, they should behave as rule minimalists in criminal law, and as rule maximalists in consumer law. A judge following this pattern would vary intrapersonally in decision-making profile, thus behaving in a way more congenial to the mental states hypothesis than to the trait-based explanation.

Hence, the hypothesis that attitudes towards rule violation judgments might be driven by traits – even if completely misguided – might turn out to be philosophically informative. Merely considering its implications and the empirical predictions that the hypothesis makes is enough to bring attention to two new and interesting profiles that add to the existing toolbox of legal decision-making, no matter whether these profiles actually obtain as traits.

Other interesting philosophical implications hinge on the trait vs. state distinction. For instance, if the results were caused by psychological states, we must investigate which precise circumstances elicit which kinds of states. This would be important both from the perspective of purely descriptive theories interested in describing what makes people more textualist or purposivist, and from the perspective of prescriptive theories siding with one of those stances, as institutional designers might explore contextual features that privilege their preferred interpretative attitude. On the other hand, if the results were caused by psychological traits, the conclusions are very different, both descriptively and prescriptively.

[115] These arguments are very underdeveloped. For instance, all cases discussed in the main text deal with rules that impose prohibitions. However, rules might impose obligations, or they might confer powers. Whether an area calls for rule minimalism, rule maximalism, or some other stance towards rules might very well depend on the kind of rule at play. In any event, hope that these inchoate arguments are sufficient to gesture towards the normative importance of considering rule maximalism and rule minimalism as autonomous decision-making stances.

5.3
The philosophical importance of experiments

Just as designing thought experiments help philosophers flesh out sophisticated analyses of concepts, designing actual experiments might also help make better philosophy. Carefully considering what conclusions can be drawn from experiments involving each set of vignettes makes salient certain conceptual imprecisions that might otherwise remain unnoticed. Similarly, counterintuitive results might shed light on new hypotheses that are of both empirical and normative import.

6

Conclusion

In this dissertation, I reviewed the available empirical data regarding the central theses of the Hart-Fuller debate. The hope was that an experimental approach could help move these important issues past the roadblock introduced by the apparently conflicting intuitions supporting each side. As a result, I reviewed 11 Studies that probed intuitions about key aspects of the disagreement between Hart and Fuller. All included, the Studies surveyed thousands of people in several different countries. What have we learned from this effort?

The goal of this conclusion is to summarize the many things we have come to know thanks to experimental investigation into the concept of rule, while listing some of the most important limitations of the existing research. Often, these limitations point towards ways in which the conclusions I drew from the data might be wrong. Future research should put these doubts to the test; thus, revisions might very well be necessary. But, as we will see, this is a sign that things are on the right path. And if experimental methods served us well with the concept of rule, it is reasonable to extend the experimental strategy that my co-authors and I employed here to other important questions in legal philosophy, such as the one about the importance of coercion to law. First, however, a brief recollection of the main arguments and ideas contained in each preceding chapter is in order.

In chapter 2, I tried to dive deeper into the content of the disagreement between Hart and Fuller over the concept of rule. I argued that the authors' positions entailed two pairs of conflicting and empirically testable theses. The first pair contrasts the Hartian idea that rules have a textual core of settled meaning that is

sometimes sufficient to determine rule violation (CSM) with the Fullerian thesis that purposes are always a necessary consideration in those judgments (SP). Both theses are empirically testable: if we can find at least one case where textual interpretation suggests that the rule was broken, but purposive interpretation calls for the opposite result, and yet most people say that the rule was violated, then CSM is true, while SP isn't. But this would tell us very little about rules. CSM is compatible both with text being treated as sufficient most of the time, and with text playing a decisive role in the rarest of occasions. Moreover, the way these theses are phrased makes it trivially easy to find empirical support for CSM, and impossibly demanding to do the same for SP. Thus, an alternate interpretation of the Hart-Fuller debate casts it as a clash between a pair of *quantitative* claims. Under that reading, Hart claimed that, in cases where textual and purposive interpretation contradict each other, text often prevails (TP), while Fuller thought that purpose would come out on top at least as often as text did (WP).

Chapter 3 turned to recent work on psychology (Turri & Blouw, 2015) and experimental philosophy (Struchiner, Hannikainen & Almeida, 2020; Almeida, Struchiner & Hannikainen, under review) for the data needed to test the four theses. Out of the gate, Studies 1 and 2 showed that CSM prevailed over SP - a result confirmed in all subsequent studies where both theses were at play. Moreover, TP also obtained in both Studies 1 and 2, with textual compliance being a much weightier predictor of rule-violation judgments than compliance with purpose. Studies 3 and 4, however, complicate the picture by showing that, under some circumstances, WP also obtains.

Study 4, in particular, raises interesting questions about the way that text and purpose interact within the concept of rule. In it, people's rule violation judgments were influenced by the presence of additional questions. When participants answered solely about rule violation, over- and underinclusion cases were both seen as borderline cases between violation and non-violation. On the other hand, when we asked participants not only about rule violation, but also whether the rule's text and purpose were violated, answers shifted, with participants treating overinclusion cases to be much clearer instances of rule violation than underinclusion cases.

What can explain this pattern of results? I have considered two rival explanations. The first one, put forth by Struchiner, Hannikainen & Almeida (2020) and Almeida, Struchiner & Hannikainen (under review), poses that rule is a dual nature concept (DNR): both text and purpose play into our judgments of rule violation. DNR explains the discrepancy between the conditions in Study 4 as either (i) a feature of the pragmatics of part-whole questions, (ii) a result of discrepancies between holistic and analytic thinking styles, or (iii) some mix of the first two explanations. This family of explanations can account for Studies 1-3, and some versions of it can explain the surprising pattern found in Study 4.

Turri & Blouw (2015), however, take issue with DNR. They suggest that the concept of rule is entirely textual, but that people sometimes answer rule violation questions in a way that is moralized. According to the authors, people behave as they do because saying that someone violated a rule entails an indirect speech act to the effect that the person is to blame[116]. This explanation (ISA) fits well with Studies 1-4. Moreover, Studies 5 and 6 gather direct evidence for it: although roughly half of the participants in Study 5 were willing to say that the rule was not violated by blameless protagonists, almost no one took that route once they could say that the rule was unintentionally violated (Study 6).

Study 7 was designed to adjudicate between these two explanations by shifting focus from rule violation to rule identity. DNR predicts that judgments about the identity of a rule should be affected by changes in purpose. On the other hand, ISA cannot account for the existence of such an effect, since there is no indirect speech act entailing blame associated with statements such as "The rule created by John is the same one used by Peter". The results, however, were not entirely straightforward: although laypeople's judgments about rule identity were affected by changes in purpose, the same effect could not be detected among lawyers.

The existence of a populational difference between lawyers and laypeople led to a reevaluation of the role of recognitional communities in legal theory. As I argued in the introduction, Hartians and Fullerians agree that certain beliefs and

[116] The flipside of that is that to say that someone didn't violate the rule would entail an indirect speech act stating that the person isn't to blame. This allows ISA to account for underinclusion cases.

attitudes of people are at least partly constitutive of law. They disagree, however, as to whose beliefs and attitudes play that role. Hart focuses on officials, while Fuller believed that the mental states of citizens also mattered. If, however, officials (assuming they are more akin to lawyers) and laypeople disagree about the concept of rule, focusing on one or the other perspective would foster very different legal theories. For instance, if laypeople usually take morally laden purposes to be just as important to rule related judgments as the rule's text, it is accurate to describe their rule-based thinking (and, hence, their legal reasoning) as intrinsically moral. However, if we focus on lawyers' textualist tendencies, it makes sense to emphasize the possibility of a disconnect between law and morality. Hence, one interesting hypothesis is that much of Hart and Fuller's more well studied downstream disagreements about the concept of law might have stemmed from an underappreciated difference in the recognitional communities each author had in mind.

Looking at the results of five experiments conducted in Brazil (Studies 1-4 and Study 7) and two in the United States (Studies 5 and 6), it seems like CSM wins over SP, and that TP and WP are true under different circumstances. But CSM, SP, TP, and WP are claims about rules *everywhere*, and they entail interesting conclusions for *general* jurisprudence, an enterprise concerned with law whenever and wherever it might arise, and not only in some particular set of jurisdictions. Hence, to say unqualifiedly that Studies 1-7 offer insight into these ambitious philosophical theses might be too hasty, especially considering that some legal philosophers, such as Tamanaha (2001) and Schauer (2015), have recently argued that this kind of generalizing project might be doomed from the very start because of the magnitude of cultural variation in law. After all, if the character of law varies widely from one instance to the other, the only claims that can be made about all instances of law are very abstract and uninteresting.

Chapter 4 reports two studies designed to test these worries. Study 8 shows that the claims of cultural variability are well founded: laypeople around the world range from extremely textualist (Poland) to decidedly purposivist (Spain). However, when we restrict our survey to lawyers (Study 9), we find that lawyers from very different cultures share surprisingly similar intuitions about rules. In the Netherlands, the U.S., Finland and Poland – countries where laypeople fall all over

the spectrum that runs from purposivism to textualism – lawyers consistently favored text over purpose. Thus, Study 9 offers an answer to sceptics about general jurisprudence: since one candidate recognitional community shares a single concept of rule across cultures, we might have interesting and consequential things to say that are true and important of law in every country. For instance: TP seems to be a pervasive feature of the way lawyers interpret rules.

Finally, Chapter 5 dealt with some new questions that arise out of the experimental investigation of the concept of rule. The first set of questions relate to the concept of purpose. Although purpose is frequently invoked by legal philosophers and despite the place it occupies in such a celebrated debate such as the one between Hart and Fuller, talk of purpose in jurisprudence is often vague. That vagueness covers clearly distinct possibilities that are fleshed out in specific debates about legal interpretation, each carrying its own set of implications for the concept of law. As Almeida, Knobe, Hannikainen & Struchiner (in prep) argue, purposes might be attributed based on (a) the original intentions of the rule maker, (b) morality, or (c) the current practice among citizens and officials. Experimental evidence suggests that each of those factors play a role in purpose attribution judgments. Future work should further specify for whom and under which circumstances each kind of purpose attribution prevails. Moreover, purposes are often implicit, but how people infer implicit purposes is something that is yet to be studied. In particular, the possibility that people might attribute multiple purposes to rules is worth considering, since the resulting proliferation of over- and underinclusion cases raises its own set of questions.

Another issue explored in Chapter 5 pertains to two possible ways to read the results of Studies 1-9. One possible explanation for the differences observed in the experiments appeals to traits: some people are purposivist, while some people are textualists. Another possibility is that the experiments capture mental states that might shift within the same persons from time to time. Consider the trait hypothesis first: it posits that people consistently interpret rules according to some pre-specified decision-making profiles. The philosophical literature about legal decision-making focuses on two of those profiles, which might be roughly characterized as hardcore textualists (those who always follow text) and passionate purposivists (those who always prefer purpose). But carefully considering the trait

hypothesis reveals that there are actually four – and not only two – decision-making profiles defined by over and underinclusion cases. Besides hardcore textualism and passionate purposivism, people may also be rule maximalists (maintaining that a rule is violated when either text or purpose is violated) and rule minimalists (a rule is only broken when both text and purpose are violated). These profiles are normatively interesting even if it turns out that they are not empirically informative, as some areas of law might call for rule maximalists, and other areas might call for rule minimalists. Moreover, if people conform to these four decision-making profiles, we can potentially say a lot of interesting things about a legal system by knowing the prevalence of each profile among the population.

This has normative implications as well. Those wishing to reform legal interpretation should care about whether traits or states offer the best explanation for people's intuitions about rules. If people behave the way they do because of personal traits, reform would be more likely to succeed if aimed at changing longstanding features of a person, for instance, by demanding that decision-makers undergo certain kinds of training. If, on the other hand, mental states explain why people rely on text or purpose at each particular case, changing contextual features related to institutional design will likely prove the most efficient way to exert influence on legal interpretation.

Overall, the existing evidence suggests a nuanced view of the concept of rule. While some very abstract things can be said of rules in every occasion (CSM), interesting claims such as TP and WP hold under different circumstances. Sometimes, people privilege text in conflicts between text and purpose, while other times they privilege purpose. The competing theories put forth to explain what determines which thesis holds (such as ISA and DNR, on the one hand, and the trait vs. states debate, on the other) are still very much alive. Even though some studies, such as Study 7, were designed to pull some of these explanations apart, the results are such that they are unable to completely discard one theory. Moreover, lawyers and laypeople have different intuitions about what matters to the concept of rule. This difference has cross-cultural implications, as laypeople's intuitions seems to vary widely with culture, while lawyers around the world converge in their textualism.

Not only are the results nuanced and compatible with multiple explanations, but many questions are still open. Chapter 5 already detailed some of them and their implications. Throughout, I sought to make the limitations of the Studies and my interpretation of the results explicit. Sometimes, this acknowledgment already contained the rough outline of what is needed to test whether my proposal stands. No doubt, future experimental jurisprudents will have the answers to many of these empirical objections. Hopefully, they will also have empirical knowledge about the success or failure of several other hidden assumptions that I am currently unaware of. Given this superior knowledge, they might look back and giggle at the naiveté of this account of the concept of rule. Maybe, they will find some fatal objection to some of the central claims I am making. For instance, if future experiments show that it is not purpose, but all-things-considered moral judgment that leads people to disagree with text-mandated judgments, a number of conclusions should be revisited.

This sort of instability marked by the possibility that future studies might substantially reshape the theoretical conclusions of an investigation is a good sign and should be welcomed. It shows that the empirical methods of experimental jurisprudence invite progress. While this might be true for most empirical sciences, it is important to note that experimental jurisprudence lies in especially shaky ground: while other fields of empirical enquiry such as psychology build upon hundreds of independently replicated findings produced by researchers at dozens of institutions, the novelty of experimental jurisprudence means that we are still dealing with a very small number of experiments coming out of a small number of universities. Moreover, these results have not yet been independently replicated. The acknowledgment of those caveats, however, shouldn't detract from the importance of experimental data. Between unsubstantiated appeals to intuition and the information gathered through Studies 1-11, we should definitely side with the latter. Especially when experimental studies suggest convincing answers to longstanding deadlocks. For instance, Hart and Fuller might have been seen – even to themselves (see Hart, 1983, p. 343) – as having such different intuitions about law as to make meaningful debate hard. However, the important difference in their answers to the question of recognitional community, coupled with populational differences in the intuitions of lawyers and laypeople, might help us explain the

divide without disregarding the views of any of the debaters[117]: maybe they were both right in their empirical presuppositions, but were nonetheless talking past each other because of the different targets of their empirical claims.

One important consideration in evaluating the results of experimental jurisprudence is the issue of ecological validity: the carefully controlled scenarios employed in the Studies I reviewed here depart in multiple and significant ways from real legal decision-making. For instance: judges never evaluate whether a single rule was violated. Usually, they have to make a determination about what ought to be done given the entire legal system. As such, they must consider how the most local rule interact with a number of other rules before saying anything about the case at hand. This means that we should be careful in extrapolating from these results to real life. Once more, this is no reason to be warry of experimental jurisprudence as a whole: future experimental work might address this issue empirically by making the scenarios more ecologically valid, for instance, by introducing complexities into the legal systems discussed in the vignettes.

While this dissertation has been primarily animated by an interest in the substantive claims made by Hart and Fuller about the concept of rule, it is also intended as an example of how experimental methods might be employed to investigate other issues in analytical jurisprudence. After all, more or less unsubstantiated appeals to intuition are abundant in other important debates. Take the debate about the importance of sanctions for the concept of law as an example. Recently, Schauer (2015) argued that contemporary analytic philosophers have seriously underestimated the importance of sanctions to law. For him, even if sanctions are not a strictly necessary feature of legal systems, they are deeply important for an adequate understanding of law. Just as the Hart-Fuller debate could be distilled into a set of opposing testable theses in chapter 2, it is quite likely that the debate between Schauer and his opponents might be similarly analyzed. If this proves to be the case, testing those theses might then help us understand the potential sources of disagreement better, just as it did in this dissertation (chapter 3). Empirical methods might also help address the issue of generality by employing

[117] Alternatively, one could argue that this is the least charitable interpretation of them all: instead of positing that one of the debaters was wrong, I am now claiming that they were both wrong in some important sense.

cross-cultural studies (chapter 4). Finally, some interesting questions that were unnoted before empirical investigation might come to light through it, informing future work in both experimental and traditional jurisprudence (chapter 5).

6.1
Practical implications

I opened this dissertation with a story about a hypothetical judge (Richard), tasked with judging whether text or purpose controlled a case (John vs. George's parents) involving same-sex marriage. In it, one of the parties (George's parents) argued for its preferred result (George's parents should inherit George's wealth) based on the rule's text[118], while the other party (John) argued for an incompatible result (John should inherit George's wealth) based on the rule's purpose. But what was introduced as a very practical scenario quickly turned into the much more abstract debate between Hart and Fuller. The entirety of the dissertation was devoted to the highly abstract theses supported by the authors and to descriptive issues regarding the evidence underlying each of them. In the end, have we learned anything of practical significance?

First, it is important to reiterate that most of the Studies I surveyed deal with whether rules were violated[119]. Crucially, none of the Studies I reviewed dealt explicitly with what *should be* done about those cases. Throughout, the assumption was that people could, at least in principle, make a distinction between whether someone violated a rule and whether they should be punished for it. All of the evidence against ISA hinges on the possibility of such a distinction. Hence, from the outset, my investigation aimed at a descriptive analysis of the concept of rule, instead of a normative theory of rule-based decision-making. However, that doesn't mean that the results are completely irrelevant for practical matters. As noted in the introduction, many of the questions that characterize normative judicial decision-making theories *presuppose* that a rule has either been violated or not. As such,

[118] "A stable union between a man and a woman, evidenced by the public engagement in continuous and durable companionship with aims to constitute a family, is recognized by the state as a family". and protected as such" – adapted from the Brazilian Civil Coded, art. 1723. The consequence at play in the hypothetical lawsuit is that those in a stable union inherit from each other in case of death.
[119] The exceptions were Study 7, dealing with rule identity, and Studies 10 and 11, dealing with purpose attribution.

advancing our understanding of which factors go into rule violation judgments might be an important preliminary to this sort of normative inquiry.

Suppose, for instance, that judge Richard wants to follow the rule-mandated result in John vs. George's parents. No matter what the rule demands, he is willing to follow it. He is, however, unsure about what the rule requires. After all, text and purpose pull into opposing directions. Could this dissertation help this puzzled version of Richard decide the case?

As I have stipulated that John, George, and Richard were all Brazilian, our data might help Richard decide by pointing out that lawyers and laypeople alike have sided with text when both text and purpose were made salient. Most Brazilians, under Richard's circumstances, would hence say that the rule-mandated result favors George's parents, even if that is immoral (but note that most Brazilians may not share Richard's stance of following the rule wherever it takes him – they may instead decide in favor of John anyway).

However, the situation would be quite different if Richard were based somewhere else. Imagine that an equally puzzled Richard is now judging a very similar case in the Netherlands. Here, the data provides less guidance: it says that lawyers would probably say that the rule-mandated result favors George's parents, but laypeople would say that the rule favors John. In order to make his decision, the Dutch version of puzzled Richard would now be confronted with the thorny issue of whose intuitions matter. On the one hand, Richard might privilege the intuitions of his peers. After all, his peers will likely read his opinion and praise or criticize it on the basis of how well *they* think he interpreted the rule. In contrast, the population at large will likely never even come to know the outcome of John v. George's parents, let alone the arguments invoked by the deciding judge. Reputational and strategic concerns are thus bound to make Richard more sensitive to the way lawyers think of rules. On the other hand, rule of law requires that rules be made public and that they should be easily understood by their addressees (see Fuller, 1969). So, it is important to apply rules in a way that is consistent with how laypeople interpret them. Which of these sets of values should prevail in adjudication? This is an entirely normative question that the Dutch version of Richard has to tackle. Empirical evidence can help illuminate which populations

are associated with each outcomes, but it cannot decide whether we should privilege the intuitions of lawyers or laypeople.

Even in Brazil, the situation is more complicated than I have let on. I said that most Brazilians under Richard's *analytic* circumstances would side with text. But most Brazilians under holistic circumstances would side with purpose. What gives the analytic judicial setting normative precedence over holistic judgments? Absent further argument, nothing.

Thus, even if the studies I have surveyed help paint a clearer picture of the concept of rule, they offer only highly conditional guidance to decision-makers. For instance, if Richard was not only puzzled in the sense of wishing to follow the rule-mandated result, but further believed that the appropriate recognitional community is made up of lawyers and that the appropriate intuitions are those elicited by analytic settings, then the fact that lawyers favor text under analytic conditions in most countries will be dispositive. However, those are many "ifs" and flipping any of them might result in a different recommendation.

This lack of empirical guidance to judges does not show a problem with the analysis. In contrast, it reveals in detailed ways the complexities of rule-based decision-making. Knowing exactly which components (such as text, purpose, and recognitional communities) are sources of such difficulties might help us make better normative sense of rule-based decision-making further down the line.

Finally, the fact that there are no clear practical implications for judges doesn't mean that there are no clear practical implications at all. In fact, from the perspective of rule makers (legal and otherwise), many of the difficulties leading to highly conditional recommendations in the case of judges vanish. If the question of which recognitional community is relevant is a thorny one from the perspective of the decision-maker, it turns into a much easier question about *addressees* from the perspective of the rule-maker. Generally, those who make rules know whether they are addressing lawyers, laypeople, accountants, economists, medical professionals, and so on. Empowered with that knowledge, they might turn to the evidence for guidance on what to privilege in the drafting process: a rule aimed at lawyers should have carefully crafted text; on the other hand, making purpose accessible might be

a more important concern when making rules for laypeople. Other features might turn out to be especially important for the intuitions of other professional populations and demographic groups.

Moreover, whilst deciding whether analytic or holistic decision-making should be preferred is hard for judges, it is easy for rule-makers to manipulate the way addressees will routinely interact with rules. For instance, procedural second-order rules might ensure that arguments pertaining to both purpose and text are considered, or they might mandate immediate appraisal of the facts, with few opportunities for discussing the interplay between text and purpose. Alternatively, the rule itself might be communicated in a way that makes text and purpose salient (maybe by spelling out both elements in language), or that bring attention only to the rule as a whole (e.g., when rules are communicated by abstract signs, such as stop signs and traffic lights). Knowing what goes into ordinary judgments of rule violation and rule identity is thus practically valuable for rule-makers that will now know the likely effects associated with the many levers they may pull.

6.2
Concluding remarks

Overall, experimental investigation of the concept of rule proved fruitful both in theoretical and practical dimensions. Theoretically, the approach helped test the claims of an enduring philosophical debate. Experimental results revealed many subtleties that are often sidelined in the armchair debate: they helped uncover the precise circumstances under which people are textualist or purposivist, as well as the importance of recognitional communities. Moreover, new hypotheses and explanations are constantly suggested by experimental work. These results, although primarily aimed at theoretical issues, also have important practical implications. They may be directly leveraged by rule-makers, and highlight important complexities involved in rule-based decision-making.

Future experimental studies will pile on. They might reinforce the findings I have reviewed, lending further credence to the claims I defended. Alternatively, they might challenge the results I reviewed in a way that forces substantial revisions to this analysis of the concept of rule. Be it as it may, further empirical evidence

will no doubt help us achieve a better understanding of the concept of rule, and, by extension, that of law.

7
References

ADLER, M. D. Popular constitutionalism and the rule of recognition: whose practice grounds U.S. law? **Northwestern University Law Review**, v. 100, n. 2, pp. 719-805, 2006.

ALEXANDER, L; PRAKASH, S. “Is That English That You’re Speaking?” Why Intention Free Interpretation is an Impossibility. **San Diego Law Review**, v. 41, i. 3, pp. 967-996, 2004.

ALEXANDER, L; SHERWIN, E. **Demystifying Legal Reasoning**. Cambridge: Cambridge University Press, 2008.

ALICKE, M. D. Culpable causation. **Journal of Personality and Social Psychology**, v. 63, n. 3, pp. 368-378, 1998.

_____. Culpable control and the psychology of blame. **Psychological Bulletin**, v. 126, n. 4, pp. 556-574, 2000.

ALMEIDA, F. N. M. R. **A nobreza togada**: elites jurídicas e a política da justiça no Brasil. PhD Dissertation (Doutorado em Ciência Política), Programa de Pós-graduação em Ciência Polícia da USP, São Paulo, 2010.

ALMEIDA, G. F. C. F.; HANNIKAINEN, I. A.; STRUCHINER, N.; BYSTRANOWSKI, P.; JANIK, B.; PRÓCHNIKI, M.; STROHMAIER, N.; BENSINGER, S.; LAAKASUO, M.; TOBIA, K.; DRANSEIKA, V.; KNEER, M.; LOPEZ, A. R.; LIEFGREEN, A.; AGUIAR, F. Rules around the world. In preparation.

ALMEIDA, G. F. C. F.; KNOBE, J.; HANNIKAINEN, I. A.; STRUCHINER, N. Purposes in law and in life. In preparation.

ALMEIDA, G. F. C. F., STRUCHINER, N; HANNIKAINEN, I. A. The concept of rule – an experimental perspective. Under review.

ALMEIDA, G. F. C. F. **Técnica Legislativa e Linguagem**. Thesis (Mestrado em Direito). Programa de Pós Graduação em Direito da PUC-Rio, Rio de Janeiro, 2017. Available at: https://www.maxwell.vrac.puc-rio.br/31028/31028.PDF.

BARAK, A. **Purposive interpretation in law**. Princeton: Princeton University Press, 2005.

BEAR, A.; KNOBE, J. Normality: part descriptive, part prescriptive. **Cognition**, v. 167, pp. 25-37, 2016.

BIX, Brian. H. L. A. Hart and the “open texture” of language. **Law and Philosophy**, v. 10, i. 1, pp. 51-72, 1991.

_____. On the dividing line between natural law theory and legal positivism. **Notre Dame Law Review**, v. 75, n. 5, pp. 1613-1624, 2000.

_____. **Law, Language and Legal Determinacy**. Oxford: Clarendon Press, 2003.

BONICA, A. Why are there so many lawyers in congress? **Legislative Studies Quarterly**, v. 45, n. 2, pp. 253-289, 2020.

BURGH, W. The work of Lon Fuller: a promising direction for jurisprudence in the twenty-first century. **University of Toronto Law Journal**, v. 64, n. 5, pp. 736-752, 2014.

BUSTAMANTE, T. Revisiting the idea of protestant interpretation. *In*: BUSTAMANTE, T.; DECAT, T. L (Eds.). **Philosophy of Law as an Integral Part of Philosophy – essays on the jurisprudence of Gerald J Postema**. Oxford: Hart Publishing, 2020, pp. 113-139.

CANE, Peter. (Org.). **The Hart-Fuller debate in the twenty-first century**. Oxford: Hart Publishing, 2010.

CASLER, K.; KELEMEN, D. Developmental continuity in teleo-functional explanation: reasoning about nature among Romanian Romani adults. **Journal of Cognition and Development**, vol. 9, issue. 3, pp. 340-362, 2008.

DAVIS, M. H. Measuring individual differences in empathy: Evidence for a multidimensional approach. **Journal of Personality and Social Psychology***,* v. 44, i. 1, pp. 113-126, 1983.

DONELSON, R.; HANNIKAINEN, I. R. *In*: LOMBROZO, T., KNOBE, J; NICHOLS, S. (Eds.) **Oxford Studies in Experimental Philosophy, Volume 3**. Oxford: Oxford University Press, 2020, pp. 6-28.

DWORKIN, R. **Taking rights seriously**. Cambridge: Harvard University Press, 1978.

FISH, S. There is no textualist position. **San Diego Law Review**, v. 42, i. 3, pp. 629-650, 2005.

FLANAGAN, B. Revisiting the contribution of literal meaning to legal meaning. **Oxford Journal of Legal Studies**, v. 30, n. 2, pp. 255-271, 2010.

FULLER, L. Positivism and fidelity to law: a reply to professor Hart. **Harvard Law Review**, v. 71, n. 4, pp. 630-672, 1958.

_____. **The morality of law**. New Haven: Yale University Press, 1969.

GARCIA, S. M.; CHEN, P.; GORDON, M. The letter versus the spirit of the law: a lay perspective on culpability. **Judgment and Decision Making**, v. 9, i. 5, pp. 479-490, 2014.

GARDNER, John. Legal positivism: 5 1/2 myths. **The American Journal of Jurisprudence**, v. 46, pp. 199-227, 2001.

_____. **Law as a leap of faith**: Essays on law in general. Oxford: Oxford University Press, 2012.

GOLDSWORTHY, J. Legislative intentions, legislative supremacy, and legal positivism. **San Diego Law Review**, v. 42, i. 2, pp. 493-518, 2005.

HANNIKAINEN, I. R.; HUDSON, N. W.; CHOPIK, W. J.; BRILEY, D. A.; DERRINGER, J. Moral migrations: desires to become more empathic predict changes in moral foundations. **Journal of Research in Personality**, v. 88, 104011, 2020.

HART, H. L. A. Positivism and the separation of law and morals. **Harvard Law Review**, v. 71, n. 4, pp. 593-629, 1958.

_____. (1983) **Essays in Jurisprudence and Philosophy**. Oxford: Clarendon Press.

_____. (1994) **The concept of law**. Oxford: Clarendon Press.

HART, H. M.; SACKS, A. M. **The legal process**: basic problems in the making and application of law. Westbury: Foundation Press, 1994.

HIMMA, K. E. Conceptual jurisprudence: an introduction to conceptual analysis and methodology in legal theory. **Revus: Journal for Constitutional Theory and Philosophy of Law**. v. 26, pp. 65-92, 2015.

HUBBARD, F. P. Power to the people: the takings clause, Hart's rule of recognition, and populist law-making. **University of Lousiville Law Review**, v. 50, pp. 87-130, 2011.

HUDSON, N. W.; FRALEY, R. C. Volitional personality trait change: can people choose to change their personality traits? **Journal of Personality and Social Psychology**, v. 109, i. 3, pp. 490-507, 2015.

HURD, Heidi. Interpretation without intentions. *In*: PAVLAKOS, G.; RODRIGUEZ-BLANCO, V. (Eds.) **Reasons and Intentions in Law and Practical Agency**. Cambridge: Cambridge University Press, 2015, pp. 52-71 doi:10.1017/CBO9781107707573.005

JOO, S.; YOUSIF, S.; KNOBE, J. Teleology beyond explanation: what factors influence teleology judgments? 2020. Available at: https://psyarxiv.com/4dtwe/. Visited October 1st, 2020.

KAHAN, D.; HOFFMAN, D.; EVANS, D.; DEVINS, N.; LUCCI, E.; CHENG, K. "Ideology" or "situation sense"? An experimental investigation of

motivated reasoning and professional judgment. **University of Pennsylvania Law Review**, v. 164, i. 2, pp. 349-439, 2016.

KELEMEN, D. The scope of teleological thinking in preschool children. **Cognition**, v. 70, i. 3, pp. 241-272, 1999.

_____. Are children “intuitive theists”? Reasoning about purpose and design in nature. **Psychological Science***,* v. 15, i. 5, pp. 295-301, 2004.

KNAPP, S.; MICHAEL, W. B. Against Theory. **Critical Inquiry**, v. 8, n. 4, pp. 723-742. 1982.

_____. _____. Not a Matter of Interpretation. **San Diego Law Review**, v. 42, i. 3, pp. 651-668, 2005.

KNEER, M.; BOURGEOIS-GIRONDE, S. Mens rea ascription, expertise and outcome effects: Professional judges surveyed. **Cognition**, v. 169, pp. 139-146, 2017.

KNEER, M.; COLAÇO, D,; ALEXANDER, J.; MACHERY, E. On second thought: reflections on the reflection defense. Available at: **<** https://www.researchgate.net/publication/325465989_On_second_thought_Reflections_on_the_reflection_defense**>**, 2020. Visited: September 23, 2020.

KNOBE, J. Experimental philosophy is cognitive science. *In*: SYSTMA, J.; BUCKWALTER, W. (Eds.) **A companion to experimental philosophy**. Hoboken: Wiley, 2016, pp. 37-52.

KNOBE, J.; SHAPIRO, S. J. Proximate cause explained: an essay in experimental jurisprudence. **University of Chicago Law Review**. Available at: **<** https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3544982#**>**. Visited: September 20, 2020. Forthcoming.

LACEY, N. Out of the “witches’ cauldron”? Reinterpreting the context and reassessing the significance of the Hart-Fuller debate. *In*: CANE, Peter. (Org.). **The Hart-Fuller debate in the twenty-first century**. Oxford: Hart Publishing, 2010, pp. 1-42.

LEVIN, H. Y. Contemporary meaning and expectations in statutory interpretation. **University of Illinois Law Review**, year 2012, pp. 1103-1158, 2012.

LOMBROZO, T.; CAREY, S. Functional explanation and the function of explanation. **Cognition**, v. 99, i. 2, pp. 167-204, 2004.

MACLEOD, J. Ordinary causation: a study in experimental statutory interpretation. **Indiana Law Journal**, v. 94, i. 3, pp. 957-1027, 2019.

MANDELBAUM, E.; RIPLEY, D. Explaining the abstract/concrete paradoxes in moral psychology: the NBAR hypothesis. **Review of Philosophy and Psychology**, vol. 3, i. 3, pp. 351-368, 2012.

MARMOR, Andrei. **Interpretation and Legal Theory**. Oxford: Hart Publishing, 2005.

NEWMAN, G. E.; BLOOM, P.; KNOBE, J. Value judgments and the true self. **Personality and Social Psychology Bulletin**, v. 40, i. 2, pp. 203-216, 2014.

NICHOLS, S.; KNOBE, J. Moral responsibility and determinism: the cognitive science of folk intuitions. *Noûs*, v. 41, pp. 663-685, 2007.

NIGRO, R. A decisão do STF sobre a união homoafetiva: uma versão pragmática da linguagem constitucional. **Revista Direito, Estado e Sociedade**, n. 41, pp. 157-183, 2013.

NYE, H. On law for crabs, angels and aliens: thought experiments and methodology in legal philosophy. Presented in the Legal Philosophy Workshop, 2020.

PHILLIPS, J., DE FREITAS, J., MOTT, C., GRUBER, J.; KNOBE, J. True happiness: The role of morality in the folk concept of happiness. ***Journal of Experimental Psychology: General***, v. 146, i. 2, pp.165-181, 2017.

PRIEL, D. Jurisprudence and necessity. **Canadian Journal of Law and Jurisprudence**, v. XX, n. 1, pp. 173-200, 2007.

ROBERTS, A. J.; WASTELL, C. A.; POLITO, V. Teleology and the intentions of supernatural agents. **Consciousness and Cognition**, v. 80, 102905, 2020.

ROSE, D. Persistence through function preservation. **Synthese***,* v. 192, pp. 97-146, 2015.

ROSE, D.; NICHOLS, S. Teleological essentialism. **Cognitive Science***,* v. 43, i. 4, e12725, 2019*.*

_____. _____. Teleological essentialism: generalized. **Cognitive Science**, v. 44, e12818, 2020.

ROSE, D., SCHAFFER, J.; TOBIA, K. Folk teleology drives persistence judgments. **Synthese**, 2018. Available at: https://doi.org/10.1007/s11229-018-01974-0.

ROSE, D.; SCHAFFER, J. Folk mereology is teleological. **Noûs**, v. 51, i. 2, pp. 238-270, 2017.

RUNDLE, K. **Forms liberate:** reclaiming the jurisprudence of Lon L Fuller. Oxford and Portland, Oregon: Hart Publishing, 2012.

SADEK, M. T.; DANTAS, H. Os bacharéis em direito na reforma do judiciário: técnicos ou curiosos? **São Paulo em Perspectiva**, v. 14, n. 2, pp. 101-111, 2000.

SCHAUER, F. Formalism. **The Yale Law Journal**, v. 97, n. 4, pp. 509-548, 1988.

_____. **Playing By The Rules**: A Philosophical Examination of Rule-Based Decision-Making in Law and in Life. Oxford: Clarendon Press, 1991.

_____. Prescriptions in three dimensions. **Iowa Law Review**, vol. 82, pp. 911-922, 1997.

_____. A critical guide to vehicles in the park. **New York University Law Review**, v. 83, n. 4, pp. 1109-1134, 2008.

_____. Positivism before Hart. **Canadian Journal of Law and Jurisprudence**, v. 24, n. 2, pp. 455-471, 2011.

_____. Hart's anti-essentialism. *In*: D'ALMEIDA, L. D.; EDWARDS, J.; DOLCETTI, A. (eds.). **Reading HLA Hart's The Concept of Law**. Oxford: Hart Publishing, 2013a, pp. 424-443.

_____. Legal realism untamed. **Texas Law Review**, v. 91, pp. 749-780, 2013b.

_____. **The Force of Law**. Cambridge: Harvard University Press, 2015.

_____. Social science and the philosophy of law. *In*: TASIOULAS, J. (ed.). **The Cambridge Companion to the Philosophy of Law**. Cambridge: Cambridge University Press, 2020, pp. 95-114.

SCHWARZ, N.; STRACK, F.; MAI, H. P. Assimilation and contrast effects in part-whole question sequences: a conversational logic analysis. **Public Opinion Quarterly**, v. 55, n. 1, pp. 3-23, 1991.

SCHWITZGEBEL, E.; CUSHMAN, F. Expertise in moral reasoning? Order effects on moral judgment in professional philosophers and non-philosophers. **Mind & Language**, v. 27, n. 2, pp. 135-153, 2012.

SHAPIRO, S. What is the internal point of view. **Fordham Law Review**, v. 75, i. 3, pp. 1157-1170, 2006.

_____. **Legality**. Cambridge: The Belknap Press of Harvard University Press, 2011.

SHECAIRA, F. P. "Interpretação jurídica e indeterminação". Thesis (dissertação de mestrado). Programa de Pós Graduação em Filosofia UFRJ, Rio de Janeiro, 2009.

_____. "A quem cabe a escolha?" Sobre a importância de distinguir os métodos alternativos de interpretação jurídica. **Revista Direito, Estado e Sociedade**, n. 36, pp. 6-23, 2010.

_____. Sources of law are not legal norms. **Ratio Juris**, vol. 28, n. 1, pp. 15-30, 2015.

SINNOTT-ARMSTRONG, W. Word meaning in legal interpretation. **San Diego Law Review**, v. 42, i. 3, pp. 465-492, 2005.

SLOCUM, Brian G. **Ordinary meaning**: a theory of the most fundamental principle of legal interpretation. Chicago: The University of Chicago Press, 2015.

SOAMES, Scott. Vagueness and the Law. *In*: MARMOR, Andrei (Ed.) **The Routledge Companion to Philosophy of Law**. New York: Routledge, 2012, pp. 95-108.

SOMMERS, R. Commonsense consent. **Yale Law Journal**, v. 129, n. 8, pp. 2232-2605, 2020.

STICH, S.; TOBIA, K. P. Experimental philosophy and the philosophical tradition. *In*: SYSTMA, J.; BUCKWALTER, W. (Eds.). **A companion to experimental philosophy**. Hoboken: Wiley, 2016, pp. 6-21.

STRUCHINER, N. **Direito e Linguagem:** uma análise da textura aberta da linguagem e sua aplicação ao direito. Rio de Janeiro: Renovar, 2002.

_____. **Para falar de regras**: o positivismo conceitual como cenário para uma investigação filosófica acerca dos casos difíceis do direito. Dissertation (tese de doutorado), Programa de Pós-graduação em Filosofia da PUC-Rio, Rio de Janeiro, 2005a.

_____. Algumas 'proposições fulcrais' acerca do direito: o debate jusnaturalismo vs. juspositivismo. *In*: MAIA, Antonio Cavalcanti; MELO, Carolina de Campos; CITTADINO, Gisele; POGREBINSCHI, Thamy. **Perspectivas atuais da filosofia do direito**. Rio de Janeiro: Editora Lumen Juris, 2005b, pp. 399-435.

STRUCHINER, Noel; HANNIKAINEN, Ivar; ALMEIDA, Guilherme da Franca Couto Fernandes. An experimental guide to vehicles in the park. **Judgment and Decision Making**, v. 15, n. 3, pp. 312-329, 2020.

STRUCHINER, Noel; ALMEIDA, Guilherme da Franca Couto Fernandes; HANNIKAINEN, Ivar. Legal decision-making and the abstract/concrete paradox. **Cognition**, v. 205, 104421, 2020.

STROHMINGER, N.; KNOBE, J.; NEWMAN, G. E. The true self: a psychological concept distinct from the self. **Perspectives on Psychological Science**, v. 12, i. 4, pp. 551-560, 2017.

TAMANAHA, B. Z. Socio-legal positivism and a general jurisprudence. **Oxford Journal of Legal Studies**, vol. 21, n. 1, pp. 1-32, 2001.

TOBIA, K. P. Personal identity and the Phineas Gage effect. **Analysis**, v. 75, i. 3, pp. 396-405, 2015.

_____. How people judge what is reasonable. **Alabama Law Review**, v. 70, i. 2, 293-359, 2018.

_____. Legal concepts and legal expertise. 2020a. Available at: < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3536564>. Visited: September 20, 2020.

_____. Law and the cognitive science of ordinary concepts. 2020b. Available at: < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3546570>. Visited: September 20, 2020.

_____. Experimental jurisprudence. 2020c. Available at: < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3680107>. Visited: September 20, 2020.

_____. Testing ordinary meaning: an experimental assessment of what dictionary definitions and linguistic usage data tell legal interpreters. **Harvard Law Review**, 133, forthcoming. Available at: < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3266082>. Visited: September 20, 2020.

TURRI, J.; BLOUW, P. Excuse validation: a study in rule-breaking. *Philosophical Studies*, v. 172, n.3, pp. 615-634, 2015.

TURRI, J. Excuse validation: a cross-cultural study. **Cognitive Science**, v. 43, i. 8, 2019.

VEGA GOMEZ, J. The Hart Fuller Debate. **Philosophy Compass**, v. 9, pp. 45-53, 2014.

WALDRON, J. Can there be a democratic jurisprudence? **Emory Law Journal**, v. 58, i. 3, pp. 675-712, 2009.

WALUCHOW, W. J. **Inclusive legal positivism**. Oxford: Clarendon Press, 1994.

_____. Legal positivism, inclusive versus exclusive. *In*: CRAIG, E. (Ed.) **Routledge Encyclopedia of Philosophy**. London: Routledge, 2001.

WILLIAMSON, T. Philosophical expertise and the burden of proof. *Metaphilosophy*, v. 42, n. 3, pp. 215-229, 2011.

_____. Armchair philosophy. *Epistemology & Philosophy of Science*, v. 56, n.2, pp. 19-25, 2019.