

ISSN 1936-5349 (print)
ISSN 1936-5357 (online)

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

EXPERIMENTAL ANALYSIS OF THE EFFECT OF STANDARDS ON COMPLIANCE AND PERFORMANCE

Constantine Boussalis
Yuval Feldman
Henry E. Smith

Published in *Regulation & Governance* (March 2017)

Discussion Paper No. 913

06/2017

Harvard Law School
Cambridge, MA 02138

This paper can be downloaded without charge from:

The Harvard John M. Olin Discussion Paper Series:
http://www.law.harvard.edu/programs/olin_center

The Social Science Research Network Electronic Paper Collection:
<http://ssrn.com/abstract=2991263>

Experimental analysis of the effect of standards on compliance and performance

Constantine Boussalis

Trinity College, Dublin, Ireland

Yuval Feldman

Faculty of Law, Bar-Ilan University, Ramat Gan, Israel

Henry E. Smith*

Harvard Law School, Cambridge, MA, USA

Abstract

Legal directives – whether laws, regulations, or contractual provisions – can be written along a spectrum of specificity, about which behavioral and legal scholarship present conflicting views. We hypothesized that the combination of specificity and monitoring promotes compliance but harms performance and trust, whereas the combination of specificity and good faith enhances both the informative goal-setting aspects of specificity and people's sense of commitment. To test these hypotheses, we used a 2x2x2 experimental design in which participants were instructed to edit a document with either general or detailed instructions, with a reference to good faith or without it, and with a review of the work or without it. Participants could engage in various levels and kinds of editing, allowing us to distinctly measure both compliance and performance. When participants require information and guidance, as in the case of editing, we found that specificity increases performance relative to the vague standard condition. We discuss the characteristics of the regulatory frameworks in which our findings are especially relevant.

Keywords: Behavioral ethics, compliance and performance, experiment, legal design, rule vs. standard, specificity.

1. Introduction

The optimal specificity of directives is one of the most basic questions in law. Most theoretical lines of research have addressed the level of discretion that should be given to adjudicators as well as to lay people who need to comply with the law. The rational choice and behavioral literatures contain a web of conflicting theories about how specific a law ought to be for it to have an optimal effect. Specificity can create a trade-off between treating individuals strictly, which ensures that minimum thresholds are met, and allowing them greater discretion, which can generate goodwill, possibly resulting in higher overall performance. In a related vein, the literature has considered what level of monitoring optimizes compliance, and when intrinsic motivation is better than monitoring at promoting compliance (Garoupa 1997, 2003). Building these subtleties into the language of the law adds yet another level of complexity to the problem.

The theoretical literature on rational choice (e.g. multi-task agency, measurement paradoxes) and behavioral issues (e.g. crowding out, justice motivation, goal setting, and resentment) has made numerous competing predictions about how these various forces interact (which we will review). Further complicating our ability to create straightforward predictions are the distinct and disparate effects of ambiguity predicted by the rational choice and behavioral literatures. For example, whereas the rational choice literature claims that ambiguity can lead to a chilling effect, the behavioral literature predicts that ambiguity can lead to under-compliance because of self-deception (about one's own compliance). As we discuss further, according to the behavioral ethics literature, people look for justification for their unethical

Correspondence: Henry E. Smith, Harvard Law School, 1575 Massachusetts Avenue, Cambridge, MA 02138. Email: hesmith@law.harvard.edu

Accepted for publication 28 September 2016.

behavior, and legal ambiguity could provide them with that justification (see [Mazar et al. 2008](#); [Feldman 2014](#); [Shalvi et al. 2015](#)).

Recognizing that the combined behavioral and rational choice bases for the effects of specificity are complex, we seek to understand which is more important for regulatory design: the effect of specificity on compliance or its effect on performance? The management literature has suggested various answers to this question in the context of organizations. We wish to extend this question to legal directives in general, where, in certain contexts, providing specific examples as to how one should behave can be helpful in encouraging individuals to follow these specific examples (compliance), but it may be harmful when we want individuals to follow the spirit of the legal directive in a broader sense (performance).

The paper offers a richer behavioral account for specificity in law, sensitive to the different functions of specificity and to the motivational mode of the individual. It is our belief that some of the conflicting views in the literature are related to an overinclusive treatment of specificity and to a limited account of the interaction between the cognitive and motivational accounts of specificity. We also replace the traditional focus on specificity as particularity and narrowness in rules with a broader consideration that includes the giving of specific examples in addition to a vaguer standard: do such examples enable people to look beyond the examples and follow the “spirit” of the vaguer standard? In addition, to achieve a more realistic understanding of the contexts in which people might move in an opposing direction, we attempt to understand specificity’s effect on the two main normative approaches: monitoring and deterrence on the one hand, and good faith and morality on the other. Based on the current literature, the combination of these two motivational aspects captures much of the relevant variation in the direction of the effect of legal specificity on compliance and performance.

The paper proceeds as follows. In Part 2, we show that the question of optimal level of detail in law is of importance for several legal doctrines. We review some of the main paradigms in this field and highlight some open areas that can be empirically investigated.

Next, we suggest that some of the reasons this question is still open in law are related to the even greater inconsistency that exists both in the rational choice and the behavioral literatures. We review and compare some of the leading theories about how people are expected to respond to specific versus vague instructions. The core of our argument has to do with the interaction between specificity and motivations for compliance. We also review the literature on intrinsic-extrinsic motivation, enforcement approaches, and the effect of measurement on behavior. Based on this review, we suggest some competing predictions about the possible effects of these three factors, individually and together.

We proceed in Part 3 to present the empirical research questions and the unique methodological design that we use to gauge the effects of specificity on behavior. We explain why the context of language enables us to compare the effect of editing instructions on a narrow notion of compliance (doing what’s required) with a broad notion of compliance (fixing errors which were not present in the specific instructions, i.e. doing more than what is required) as well as of a third category of errors which required extra effort or went beyond reasonable expectation, which we call performance (e.g. content-related fixes unrelated to problems in the language of the document participants were required to fix).¹

To create an empirical paradigm to allow for a comparison of the two types of behaviors – compliance and performance – we have focused on a particular genre of legal specificity, which we refer to as *example-giving*. While people might usually refer to driving 55 miles per hour as the classical prototype of specificity, we believe the dilemma of whether regulators should provide laws with detailed examples as to the meaning of the law is highly important for many legal doctrines. Part of this type of specificity has recently received analytical treatment by Parchomovsky and Stein (2015) who define example-giving as a “catalog” approach. According to their approach, which focuses mainly on interaction between the legislative and judicial branches, in many areas of law there is a middle hybrid approach that employs vague standards on the one hand and detailed examples on the other, which is aimed at helping people understand the rationale of the specific law.² Indeed, in reality, the situation is even broader: specific legal directives (laws, regulations, or contracts) do not provide any general standard regarding the overall purpose of the law, regulation, or contract, aside from the specific instructions. Hence, we believe that understanding whether specific examples aimed at itemizing a legal standard are beneficial to the people who need to follow them is important for numerous private law doctrines, as well as many types of commercial contracts.

In Part 4, following a presentation of the findings of the study on the effect of example-giving on compliance and behavior under conditions of either good faith or monitoring, we examine the implications of the findings for legal contexts in which some level of specificity needs to be chosen and we elaborate on the limitations of the study.

2. Existing approaches to specificity

2.1. Accounts of specificity in law and economics

In legal scholarship, much of the debate on the optimal specificity of law comes from the “rules versus standards” paradigm, where standards are more vague than rules. Kaplow’s (1992) seminal paper on this topic and many follow-up studies translated the notion of “optimal specificity” into “optimization of information costs.” Kaplow includes various costs and benefits of rules versus standards in his model, but broadly speaking, rules are more costly to create and cheaper to enforce.

Some economic analyses point to the benefits of vagueness, but these benefits are often the flip side of costs. Thus, vagueness can smooth out the liability function, which may reduce the cost of errors in selecting the point at which the sanction sets in (Cooter 1984; Craswell & Calfee 1986). Contracting parties often include vague terms in their contracts, which can serve as a commitment device that increases the cost of litigation (Choi & Triantis 2010). An additional perspective comes from the multi-tasking paradigm that focuses on the problems that occur when aspects of someone’s work are easier to monitor than others. According to the rational choice prediction, the agent focuses most of her work on the tasks for which she can be given an incentive (Holmstrom & Milgrom 1991).

An additional line of research in law and economics focuses on the chilling effect, a form of over-compliance. Ferguson and Peters (2000, p. 25) argue “the optimal amount of vagueness in a rule strikes a balance between the costs of loopholes, the chilling effect on economic activity, and the inefficiency created in the legal system.” Economic analyses focus on how risk-neutral agents might over-comply with a vague law if the law is built around the optimal standard, albeit vaguely. For example, Logue (2007) and others have noted that risk-averse taxpayers over-comply with an uncertain legal rule coupled with high penalties. Craswell and Calfee (1986) showed that risk-neutral agents would be subject to two behavioral effects of legal uncertainty: uncertainty would reduce deterrence because of the prospect of escaping liability wrongfully (a false negative), but in a variety of contexts, this effect would be dominated by a tendency toward over-compliance, which has been recognized, for example, in tax law. Under a vague standard, uncertainty can cause damages to rise more quickly than social harm around the optimal point, leading the actor to reduce the expected liability by inefficiently over-complying.

2.2. The behavioral perspective

From a behavioral perspective, compliance with rules that do and do not include detailed examples is not solely a matter of decision under risk, but rather it encompasses a whole array of cognitive (e.g. attention, information) and motivational (e.g. resentment, trust and commitment) factors. Hence, the behavioral account of specificity supplements the rational choice perspective in a few ways: it provides a richer and more nuanced view of the different functions served by specificity and, in addition, it enables a better understanding of the interaction between specificity and the motivation people have toward the behavior. Our study demonstrates that, by taking this nuanced approach to specificity, we can improve the predictability of the effect of specificity on behavior in different contexts.

On the motivational side, behavioral research focuses on aspects such as work morale and resentment that may lead subjects to view specificity and over-detailed rules as a sign of mistrust and as a signal to extrinsic rather than intrinsic motivation (i.e. crowding out). On the positive side, in the literature on motivated reasoning and self-concept maintenance, specificity might reduce the option for self-serving biases. The cognitive literature also points in multiple directions. In a negative cognitive effect, specificity causes subjects to pay too much attention to detailed instructions and to succumb to an availability bias, leading to unintended ignorance of other unspecified aspects. In contrast, the positive cognitive side of the specificity literature, such as that on goal-setting, suggests that specificity may increase work focus better than “do your best” (see also Coglianese *et al.* 2003).

An additional way in which the behavioral literature contributes to the rational choice discussion is by emphasizing the distinction between compliance and performance. Elsewhere we have demonstrated the importance of this dichotomy for legal theory (see Feldman & Smith 2014). While in compliance we focus on people doing exactly as required, in performance we measure whether people make an extra effort to fulfill the spirit of the request (see Garcia *et al.* 2014). There are also various intermediate levels of behaviors, which the literature sometimes terms “beyond compliance” or extra role behavior (see, e.g. Kim & Mauborgne 1996).

In the following paragraphs, we develop these conflicting views in the literature in a clearer and more analytical way.

2.2.1. Behavioral advantages of specificity. The main psychological theory that challenges the negative views of specificity in the behavioral literature is “goal setting.” The theory, which is highly influential in the management literature, claims (and data have shown) that challenging goals lead to higher performance than vague ones, such as “do your best.” For example, [Locke *et al.* \(1988\)](#) reported 24 field experiments that found that individuals with specific, challenging goals either outperformed others or outperformed their own past performance when they had been instructed to simply “do their best.” [Latham and Yukl \(1975\)](#) reviewed earlier studies of performance and similarly concluded that when people are given specific, complex goals they perform better than when being asked to simply “do their best.” A classic study conducted by [Seijts and Latham \(2001\)](#) compared giving people “do your best” instructions versus assigning distal and proximal goals, and found that the combination of goals led to better performance than the “do your best” approach. However, the “do your best” approach produced better performance than the assignment of distal goals only. Therefore, this line of research seems to suggest that when specificity helps to provide clear instructions and when it increases feelings of self-efficacy, it is superior to ambiguous standards or instructions. However, it should be noted that even the goal-setting paradigm recognizes contexts in which specific instructions can cause underperformance compared with ambiguous ones. For example, [Zhou and Shalley \(2003, p. 179\)](#) have argued “it is possible that goals that direct individuals’ attention toward completing more task units would simultaneously direct their attention away from coming up with creative ideas about their work.” p. 179.

As suggested earlier, when considering the behavioral advantages of specificity, we must account not only for cognitive but also motivational effects. Indeed, behavioral research, particularly in behavioral ethics, has found a correlation between a preference for ambiguity and a desire to justify one’s questionable behavior. [Haisley and Weber \(2010\)](#), for example, found that people prefer ambiguous risks when such ambiguity allows them to justify their unfair behavior, and [Dana *et al.* \(2007\)](#) found that people are less generous in situations in which they can appeal to moral ambiguity to explain their selfish behavior. Similarly, [Hsee \(1995\)](#) found evidence that people make choices that satisfy their own preferences at the cost of not completing an assigned goal if they can exploit existing ambiguity about what decision could be considered to achieve the assigned goal (see also [Ayal & Gino 2011](#)).

Proposition 1A³: Based on theories of goal setting and self deception, we propose that specificity will increase the likelihood of compliance with specific legal instructions, at least in a narrow sense, relative to an ambiguous instruction that doesn’t provide any specific example.

2.2.2. Behavioral disadvantages of specificity. The jumping off point for the behavioral literature is the rational choice literature on how various methods of providing incentives will effect compliance through measurement, monitoring, and sorting. For example, a classic study by [Lazear \(2000\)](#) shows that the piece-rate approach, in which people are evaluated based on how many units they produce, results in higher performance than a fixed wage approach. These studies also argue that employees react only to measurable metrics, especially in situations where it is difficult to determine who does what. By contrast, aspects of the work that are not easily measured suffer from poor performance. In the measurement paradox and the related multi-tasking effects problem, over time the accuracy of measurement decreases as people concentrate their effort strictly on the measured components of an activity, resulting in a decline in the overall quality of their performance ([Holmstrom & Milgrom 1991](#); [Prendergast 1999](#)). The behavioral literature likewise focuses on hard to measure dimensions of compliance and performance. Thus, people’s intrinsic motivation to perform well is crowded out by the relationship between performance, measurement, and payment ([Bowles 2008](#)). Therefore, specificity combined with monitoring that focuses only on given measurable components (the letter of the law) seems to produce a straightforward effect of crowding out intrinsic motivation and decreasing overall performance.

[Chou *et al.* \(2010\)](#) have recently applied this concept to the area of contract specificity and have shown how feelings of distrust, triggered by an overly specific contract, can lead people to low performance in a long-term contract and serve as an obstacle between the two sides in long-term relationships (see also [Poppo & Zenger 2002](#)). The authors rely

on a long list of negative psychological mechanisms, mostly related to motivation, which could be triggered when the parties create a contract. Here the overall prediction is that, with a less specific contract, employees will exert greater effort, strive for greater efficiency, and act in a more trustworthy way. The authors support these theoretical predictions with a series of experiments in which participants who were given a highly specific contract were less likely to perform well relative to those who were given a less specific contract. For example, the authors compared the effect of “notify within one hour” with that of “notify as soon as possible” as part of a comparison between specific and less specific contract terms. The specific term – “within an hour” – is not just clear and informative, it is also framed in a way that completely limits one’s choice, especially relative to what has been defined as a less specified condition – “as soon as possible” – which is not just flexible but also gives the other party a very strong signal of respect. We employ a different approach in this paper to examine the effect of specific rules. We do not constrain people’s choices, but rather we add some detailed examples to the vague standard (see Parchomovsky & Stein 2015).

Possibly the most advanced line of research involving the inadvertent cognitive effects of specificity concerns the debate around the efficacy of using checklists to ensure compliance (Gibbons & Henderson 2013). Most of the literature deals with the checklist as a way of reducing human errors, especially in aviation and hospitals, by helping the staff maintain self-control in stressful situations. The checklist has gained its fame mainly in the cockpit, where pilots found their routine missions too long and complex to remember. Although the efficiency of the checklist is no longer contested, the way to build and use checklists is still debated. Using a checklist on a daily basis can result in overuse, leading to low performance by making tasks automatic. Depending on its nature, a checklist can impair quality, reduce the expediency of services, and interfere with professional judgment and objective decisionmaking.

Proposition 1B: Based on theories of crowding out, resentment, and deviation of attention, we propose that specificity will reduce the likelihood that people engage in a broad sense of compliance (less easily measured margins), as well as performance.

2.3. The relationship between specificity and intrinsic/extrinsic motivation

As hinted above, a way to possibly better understand the relationship between the cognitive and motivational effects of specificity on both compliance and performance is related to its interaction with the motivational modes of those processing the language of the legal instrument. Many of the conflicting studies discussed above, for and against increasing the level of specificity of legal standards, fail to account for the role of motivation in a systematic way.

2.3.1. The importance of intrinsic motivation. Research on why people obey the law, starting with that of Tyler (Tyler 2006), demonstrates the importance of non-instrumental motivation, such as individuals’ desire for fairness, in accounting for compliance and performance. Frank (1988) recognized moral motivation as a force that encourages people to defy narrow versions of rational choice theory, making decisions that are either neutral or contrary to material self-interest. Much of the original behavioral economics literature examined how concepts of fairness may be employed to encourage individuals to overcome their own self-interest (see also Shavell 2002). Other scholars have focused on the role of morality and, in particular, on the content of the law rather than on procedure as a mechanism for increasing compliance (Tyler & Darley 2000). The unique contribution of morality to legal compliance has been demonstrated in various legal contexts, even in areas that are usually viewed as economic, such as taxation (Wenzel 2005). More broadly, economists argue that the law can cultivate social norms, which in turn serve as intrinsic motivation for compliance with the law (Cooter 2000). Cooter’s work demonstrates the growing recognition among law and economics scholars of the superiority of compliance triggered by intrinsic motivation, as a result of the lack of reliance on monitoring, enforcement, and stability (see also Bardach 1989 and Feldman 2011 for a review). More recent work in the context of ethical decisionmaking (e.g. Aquino *et al.* 2009) has also suggested that it is possible to increase ethical behavior through various types of exposure to ethical text (Welsh & Ordóñez 2014). Thus, this literature suggests that reminding people of moral norms will increase their motivation to behave morally and hence, reminding people that they need to behave in “good faith” is expected to positively affect both their compliance and performance.

2.3.2. Interplay between specificity, monitoring, and compliance motivation. From a policymaking perspective, the relevant question is not merely whether intrinsic motivation can increase compliance or performance, but whether and

when intrinsic motivation outperforms extrinsic motivation in securing desirable behavior. In contexts in which intrinsic motivation is more successful in securing compliance, policymakers must make sure that the standard extrinsic motivators provided by law (e.g. sanctions and incentives) do not undercut intrinsic motivations. This issue has received attention mostly in the literature on crowding-out effects. We review some of the findings of this research, including its shortcomings, and then examine the conflicting literature that supports the idea that extrinsic motivation can serve to increase intrinsic motivation.

As noted earlier, extrinsic motivation is generally thought to undermine intrinsic motivation. Fehr and Falk (2002) argued that when people attribute their behavior to external rewards, they discount their moral incentives for their behavior, thereby lowering the apparent effect of intrinsic motivation. For example, paying people to donate blood causes donors to view the donation as a transaction rather than a charitable act, eroding altruistic blood donation. Similarly, in a series of lab-based experiments, Deci (1971) found that tangible rewards undermine intrinsic motivation for a range of activities. In another study, Deci *et al.* (1999, p. 653) argued, “tangible rewards tend to have a substantially negative effect on intrinsic motivation.” In a similar vein, Marshall and Harrison’s (2005, p. 653) work suggests that the use of incentives can damage self-esteem, resulting in the perception that professionalism is no longer valued.

Likewise, with sanctions, there appears to be a trade-off between intrinsic motivation and compliance. In the area of taxation, increasing the power aspect of authorities through devices like fines and higher auditing rates can reduce the trust of taxpayers in the taxing authority, and can reduce overall compliance (see, e.g. Wenzel 2005; Kirchler *et al.* 2008). Under a broad set of circumstances, gentle rule enforcement (between no enforcement and total enforcement) can lead to more effective compliance (Erev *et al.* 2010).

Yet another nuance of this theory suggests that the effect of incentives is not linear, but rather that “intermediate” payouts have a disproportionately high crowding-out effect compared with low or high payout levels. Tenbrunsel and Messick (1999) argued that a weak system of sanctions produces worse results than having no sanctions at all. By introducing a sanctioning system, the principal changes the evaluation of the problem from an ethical dilemma to a business decision, consequently shifting individual considerations away from ethics and toward self-interest.⁴

In contrast to the above research, some scholars in the accounting literature suggest that, over time, a control system and sanctions can increase the level of cooperation among partners and increase long-term trust (Coletti *et al.* 2005). These possible influences raise concerns about whether such effects would be manifested also in the context of law. Work by Feldman and Tyler (2012) offers initial support for the suggestion that imposing a law on an existing social practice would enhance, rather than undermine, the effectiveness of the social practice.

Specificity may also affect the relationship between agent and principal, making the agent resentful toward the principal as a result of fear or lack of trust. Lack of trust may be another route by which specificity decreases intrinsic motivation and the resulting performance on a given task. We have already noted that a series of studies by Chou *et al.* (2010) examined how people perceive specific contracts, and their results suggest that specificity crowds out intrinsic motivation through perceptions of mistrust (see also results by Fehr and Gächter 2000). This conclusion is also supported by an earlier study by Tenbrunsel and Messick (1999), who argued more broadly that strict enforcement, sanctions, and specificity could harm cooperation. Dickinson and Villeval (2008) maintained that when principals engage in costly monitoring, agents react to the resulting discipline of monitoring by increasing effort. However, they also reported that intrinsic motivation is crowded out when monitoring exceeds a certain threshold. Overall, they tend to agree with the views of Osterloh and Frey (2013, p.9) that “A job providing decision latitude enhances employees’ self-determination and thereby strengthens interest and pride in the job.” Their overall conclusion is that monitoring negatively affects the principal–agent relationship rather than crowding out intrinsic motivation.

A similar view is supported by Falk and Kosfeld (2006), who demonstrated the effect of specificity on the principal–agent relationship in experiments in which the principal could either let an agent decide a production amount (ambiguous instruction) or the principal could set a lower limit for production (specific instruction). When a specific lower limit was set, agents produced less than when the principal left the production levels to the agent’s discretion. In post hoc questioning, agents stated that they saw the specific lower limit as a signal of distrust and therefore behaved less cooperatively. Specific instructions also give individuals less room for discretion, creating a situation in which the individual must constantly look for external instructions.

Proposition 2A: Based on the research on crowding out and signals of mistrust, a combination of specificity with high external controls will increase compliance but will reduce performance.

Proposition 2B: Along the same lines, a combination of specificity with a focus on morality and good faith will mitigate the negative motivational aspects of specificity and will increase both compliance and performance.

3. Experimental method

To determine how to bring together the approaches to the optimal level of specificity, in the next section we explore the effect of specificity experimentally, in a way that allows for a comprehensive examination of competing theories about how specificity affects both compliance and performance. We explain the methodological approach we follow in the present paper to further our understanding of how the type of specificity employed in the law interacts with some of the main relevant enforcement mechanisms related to intrinsic and extrinsic motivation: monitoring and good faith.

The combination of specificity, monitoring, and good faith in a 2x2x2 design, using a variety of incentive-compatible behaviors enables us to explore a spectrum of effects of legal specificity across a range of contexts. Understanding how the use of specific instructions affects behavior both alone and in tandem with monitoring and a good faith requirement, allows us to develop a coherent understanding of how to optimize the language of the law, regulation, and contract in order to encourage compliance and performance. Furthermore, the behavioral dependent variable chosen for this project is language editing, which gives us the freedom to vary the level of specificity. Although we assume that most participants have some notion of what editing means, the task of editing still makes it possible to emphasize some aspects in greater detail without undermining the overall meaning of the activity. It is also an area in which some objective criteria for evaluation exist, but levels of performance can vary, permitting us to determine the extent to which participants were merely doing the minimum or trying their best.

3.1. Design

In the study, we explore the relationship between specificity, good faith, and monitoring on compliance and performance. Specifically, we seek to determine the effect of contract-style specificity on editing effort by providing detailed examples to only half of the participants. Editing is a good experimental stand-in for the type of behavior governed by contracts. Contracts can cover a large swath of potential behavior, and editing falls between the extremes of open-ended creativity and purely rote or one-shot acts (such as the transfer of a good). It is in this middle range of behavior that the factors we are interested in are most likely to manifest themselves and where legal design is likely to be least straightforward. This choice is in line with previous experimental studies of contracting behavior (see, e.g. Brooks *et al.* 2014 for a study of quality thresholds using a task involving counting the “1” digits in 200 figures between 1 and 9, designed to hone in on effort). As suggested above, while we are interested in understanding the effect of increasing or decreasing specificity on people’s compliance and performance, we are not comparing the two extreme conditions as some of the other studies on specificity have suggested (Chou *et al.* 2010), but rather the effect of a general instruction versus a condition where we add some specific examples (out of many other categories) of what the vague standard means.

We operationalize compliance and performance as the number of *attempted* corrections of a text document that contains nine types of errors: verb tense, punctuation, subject-verb agreement, missing words, word meaning, pronoun usage, article usage, singular/plural, and apostrophe usage. There were a total of 54 mistakes in the 1,592-word document that was distributed for editing. The experiment was conducted online in October 2012, with a sample of 339 respondents who were recruited from Amazon.com’s Mechanical Turk (MTurk), a crowd-sourcing Internet labor market.⁵ Respondents were paid \$1.50 for participating in the study (58% female; mean age = 32.48, standard deviation [SD] = 11.23).⁶

Respondents were first shown a page with information about the study and a statement of informed consent. After a respondent opted into the study, he or she was then randomly assigned to one of eight experimental conditions that primed respondents with different combinations of specificity, good faith, and monitoring (2x2x2). Prior to editing the document, subjects were shown text prompts that differed based upon their treatment assignment (see Appendix 6.3 for exact wording of the prompts). Following is an overview of the treatment manipulations.

3.1.1. Specificity manipulation. Respondents who were exposed to the specificity condition were informed that while editing the document, “the typical types of errors to which [they] should pay particular attention . . . include, but are not limited to, mistakes in punctuation and verb tense.” Respondents were then provided with examples of proper punctuation and verb tense usage, followed by a short multiple-choice quiz. Those who were not assigned to the specificity condition (ambiguous condition) were told that they were to carefully edit a text document, but that they were not being given “specific examples of what kind of errors [they] should focus on.” The specificity manipulation was intended to highlight a subset of specific types of errors, while also reminding the respondent that there were other error types not mentioned. The expectation was that participants who were given instructions with specific error type examples would focus more on these “mentioned” examples of error types when correcting the document. The seven error types not mentioned were either *grammatical*, which require a “reasonable” amount of attention to detect, or *content-related*, which require a “more than reasonable” level of focus. We expected that the specificity condition would increase the level of editing of the mentioned error types and reduce editing performance for non-mentioned types of errors.

3.1.2. Good faith manipulation. Participants assigned to the good faith treatment were instructed that they were to act in good faith while editing the text document, meaning that they “should act in a genuine, sincere, fair, and honest manner while participating.” Furthermore, those in the good faith condition were also asked to write a few sentences on what good faith means to them, as well as what it means to edit a document in good faith. Subjects who were not assigned to the good faith treatment were not shown any information regarding good faith. We expected that subjects with specific instructions and who were assigned to the good faith conditions would be less likely to ignore non-mentioned error types.

3.1.3. Monitoring manipulation. Subjects assigned to the monitoring condition were instructed, “satisfactory execution of the editing will be rewarded with an extra \$1.00 bonus payment.” Before making this payment, however, study administrators “will carefully review [subjects’] performance as well as how closely [they] followed our instructions.” Moreover, respondents in the monitoring condition were asked to estimate the number of minutes they believed the researchers would invest in verifying that the instructions were followed satisfactorily before payment of the bonus. Subjects who were not assigned to the monitoring manipulation did not receive any information with respect to a potential review of work by administrators and were also not asked to think about the amount of time the researchers would invest in quality control of the edits. Participants with specific instructions of the monitoring condition should be more likely to edit only what is required (mentioned error types).

3.1.4. Dependent variable. Summarizing the dependent variable employed in this study, we operationalized compliance and performance with the number of attempted copyediting corrections over nine distinct types of errors. Because of the nature of how the specificity treatment was framed, we can classify these error types into three categories of compliance and performance: *narrow compliance*, *broad compliance*, and *performance*. Respondents who were given specific instructions were informed of the presence of two types of errors (punctuation and verb tense). As discussed above, we expected that a minimal level of compliance with the instructions (narrow compliance) would involve editing these explicitly mentioned error types. The unmentioned error types fall under two categories. The first were grammatical errors that objectively require low levels of effort in order to detect, such as singular/plural mistakes. Respondents given specific instructions and who edited these types of errors displayed broad compliance: effort that went beyond the explicit instructions but which was simultaneously objectively easier to monitor. The second category of unmentioned error types encompassed the notion of performance. These content-related errors, such as word meaning, required a closer reading of the text and, as such, detection of these errors signaled that the respondent went “above and beyond” the explicit instructions provided.

To minimize measurement error and also economize resources, we employed an automated approach to evaluate editing effort. Respondent editing performance was measured using an algorithm, relying on a set of regular expressions, which searched each edited document, found and flagged all the instances in which a change had been made to incorrect text, and tallied the number of attempted corrections made by the respondent for each type of error (e.g. punctuation, verb tense, subject-verb agreement, etc.).

3.2. Methodology

Because we assigned individuals randomly to the various treatment conditions, we made the concepts of monitoring, good faith, and specificity salient before the participants completed the editing task. Thus, we were able to directly compare whether the number of corrected errors differed across these conditions. The data were analyzed using a three-factor analysis of variance model, with the specificity, good faith, and monitoring treatment conditions as the factors. As there was only one error in the categories of article usage, singular/plural, and apostrophe usage, we used logistic regression to estimate the effects of the experimental conditions. A statistically significant positive ($P < 0.01$) main effect of specificity on post-treatment English knowledge performance led us to also include knowledge of English as a control in all models of the study.

3.3. Results

3.3.1. Narrow compliance. In the instructions provided to subjects given the specificity treatment, there was a mention that, among other types of errors, verb tense and punctuation errors were present in the document. The statistical results of the effect of the treatments on editing effort for these errors approximating narrow compliance are as follows.

With regard to verb tense errors, the specificity treatment significantly increased the editing effort. As shown in Table 1, we found that respondents exposed to the specificity treatment corrected more verb tense errors ($M = 4.945$, standard error [SE] = 0.236) than those in the ambiguous treatment ($M = 3.90$, SE = 0.248), $F(1, 330) = 9.211$, $P < 0.01$.⁷ We did not determine any other treatment effect on editing effort for verb tense errors.

Moving on to the second error type for narrow compliance, punctuation errors, as shown in Table 2, the interaction of monitoring and specificity was statistically significant, $F(1, 330) = 4.478$, $P < 0.05$, with respect to the number of punctuation error corrections. Respondents exposed to the specificity condition without monitoring performed better ($M = 5.215$, SE = 0.229) than those assigned to the ambiguous treatment without monitoring ($M = 4.481$, SE = 0.256). Specificity produced no significant difference in performance among those in the monitoring condition.

The results presented in Tables 1 and 2 conform to our theoretical expectation that respondents in the specificity treatment condition would correct significantly more “mentioned” error types, which correspond with narrow compliance.

3.3.2. Broad compliance. A second category of editing effort in the study was broad compliance, which is approximated by five error types: subject-verb agreement, pronoun usage, article usage, singular/plural, and apostrophe usage. For the most part, we did not find any treatment effects for these error types. The only exception was subject-verb agreement, which is displayed in Table 3. Specifically, we found a statistically significant and positive effect, $F(1, 330) = 11.684$, $P < 0.001$, of specificity on the level of subject-verb agreement corrections, with the specificity group showing a higher ($M = 7.095$, SE = 0.253) number of corrections than the ambiguity group ($M = 5.832$, SE = 0.266).

Table 1 Three-way analysis of variance for attempted verb tense corrections

	Partial SS	df	MS	F	P
Model	383.430	8	47.929	5.000	0.000
Monitoring	0.075	1	0.075	0.008	0.930
Good Faith	0.001	1	0.001	0.000	0.992
Monitoring*Good Faith	13.101	1	13.101	1.367	0.243
Specificity	88.306	1	88.306	9.211	0.003
Monitoring*Specificity	0.080	1	0.080	0.008	0.927
Good Faith*Specificity	2.771	1	2.771	0.289	0.591
Monitoring*Good Faith*Specificity	7.870	1	7.870	0.821	0.366
English	215.171	1	215.171	22.445	0.000
Residual	3163.549	330	9.587		

Table 2 Three-way analysis of variance for attempted punctuation corrections

	Partial SS	df	MS	F	P
Model	237.660	8	29.708	6.298	0.000
Monitoring	1.448	1	1.448	0.307	0.580
Good Faith	2.044	1	2.044	0.433	0.511
Monitoring*Good Faith	0.192	1	0.192	0.041	0.840
Specificity	4.217	1	4.217	0.894	0.345
Monitoring*Specificity	21.122	1	21.122	4.478	0.035
Good Faith*Specificity	0.271	1	0.271	0.057	0.811
Monitoring*Good Faith*Specificity	1.548	1	1.548	0.328	0.567
English	195.069	1	195.069	41.352	0.000
Residual	1556.699	330	4.717		

Table 3 Three-way analysis of variance for attempted subject-verb corrections

	Partial SS	df	MS	F	P
Model	543.592	8	67.949	6.169	0.000
Monitoring	16.366	1	16.366	1.486	0.224
Good Faith	0.906	1	0.906	0.082	0.774
Monitoring*Good Faith	2.308	1	2.308	0.209	0.647
Specificity	128.703	1	128.703	11.684	0.001
Monitoring*Specificity	0.011	1	0.011	0.001	0.975
Good Faith*Specificity	0.052	1	0.052	0.005	0.945
Monitoring*Good Faith*Specificity	15.042	1	15.042	1.366	0.243
English	283.956	1	283.956	25.778	0.000
Residual	3635.069	330	11.015		

3.3.3. *Performance.* The final editing category was content-related errors where editing effort would constitute performance; specifically, these errors are missing words and word meaning. Beginning with the missing words error type, the results of the analysis are shown in Table 4. We found two statistically significant main effects. First, the respondents exposed to the monitoring treatment made fewer corrections ($M = 2.391$, $SE = 0.152$) than those without the monitoring condition ($M = 2.883$, $SE = 0.157$), $F(1, 330) = 5.103$, $P < 0.05$. Second, respondents in the specificity treatment group made more missing word corrections ($M = 2.855$, $SE = 0.151$) than those exposed to ambiguous instructions ($M = 2.419$, $SE = 0.159$), $F(1, 330) = 3.894$, $P < 0.05$.

We also found a significant interaction effect between monitoring and good faith at the 10 percent error level for the missing words model. We found that respondents who were not exposed to the good faith treatment but were treated

Table 4 Three-way analysis of variance for attempted missing word corrections

	Partial SS	df	MS	F	P
Model	123.573	8	15.447	3.927	0.000
Monitoring	20.074	1	20.074	5.103	0.025
Good Faith	0.037	1	0.037	0.009	0.923
Monitoring*Good Faith	12.295	1	12.295	3.125	0.078
Specificity	15.317	1	15.317	3.894	0.049
Monitoring*Specificity	3.571	1	3.571	0.908	0.341
Good Faith*Specificity	4.887	1	4.887	1.242	0.266
Monitoring*Good Faith*Specificity	3.035	1	3.035	0.772	0.380
English	49.834	1	49.834	12.669	0.000
Residual	1298.096	330	3.934		

Table 5 Three-way analysis of variance for attempted word meaning corrections

	Partial SS	df	MS	F	P
Model	80.520	8	10.065	3.998	0.000
Monitoring	1.490	1	1.490	0.592	0.442
Good Faith	1.100	1	1.100	0.437	0.509
Monitoring*Good Faith	5.838	1	5.838	2.319	0.129
Specificity	12.798	1	12.798	5.084	0.025
Monitoring*Specificity	3.738	1	3.738	1.485	0.224
Good Faith*Specificity	0.438	1	0.438	0.174	0.677
Monitoring*Good Faith*Specificity	0.357	1	0.357	0.142	0.707
English	44.628	1	44.628	17.729	0.000
Residual	830.707	330	2.517		

with the monitoring condition performed worse ($M = 2.209$, $SE = 0.201$) than those who were neither in the good faith nor in the monitoring conditions ($M = 3.086$, $SE = 0.233$). Monitoring had no significant effect on performance among the participants assigned to the good faith condition.

The results for the treatment effects on editing effort for word meaning errors are shown in Table 5. We found a positive and statistically significant effect, $F(1, 330) = 5.084$, $P < 0.05$, of specificity on the number of word meaning corrections: respondents presented with specific instructions had an average of 2.552 ($SE = 0.121$) attempted corrections while those given ambiguous instructions prior to the editing task had an average of 2.154 ($SE = 0.127$) attempted corrections.

Thus, when it comes to non-required tasks, as well as those that require extra attention, specificity did not have a similar effect to that of monitoring and led people to perform better in those tasks.

3.4. Summary of results

The empirical results suggest a nuanced role of specificity on editing effort. In line with Proposition 1A, we found that respondents in the specificity condition who were provided with a list of mentioned error categories were more likely to attempt to correct these types of errors than respondents who were given ambiguous instructions. However, contrary to Proposition 1B, we found that subjects in the specificity condition were also more likely to correct many of the unmentioned types of errors. The main effect results of specificity suggest that there is some type of spillover effect, given that respondents with the specificity treatment not only performed better with respect to “mentioned” (narrow compliance) error types but also with respect to “unmentioned” error types. Of the “unmentioned” errors, we found a significant main effect on all content-related types, which approximate performance (missing words and word meaning), as well as one of five grammatical types (subject-verb agreement), which conform with broad compliance.

We attribute the spillover effect to an increase in the level of attention and not necessarily to heightened intrinsic motivation. We believe that by offering examples of error types to these respondents, they were more careful and thus more likely to identify errors of all types. Respondents exposed to the specificity treatment were more likely to attempt to correct all of the content-related error types (high effort) and some grammatical-related types (lower effort). These results may reflect more of the instructional aspect of specificity than notions of morality/specificity. That is, for certain tasks, such as copyediting, the provision of specific instructions may not necessarily affect the intrinsic motivation of respondents. In certain contexts, specific instructions may indeed be sought after and welcomed by agents who would otherwise perceive the task as prohibitively ambiguous.

With respect to our theoretical expectations laid out in Propositions 2A and 2B, we find little evidence. As a reminder, Proposition 2A states that the combination of specific instructions and monitoring would lead to increased compliance but reduced performance. Of the nine error types, we found a statistically significant interaction effect of these two conditions only for punctuation errors, which is an error type mentioned in the specific instructions. However, we found that among those given specific instructions, respondents who were *not* given a monitoring treatment did *better* at detecting punctuation errors than subjects who were given a monitoring treatment. To further illustrate the mixed results, we found no significant interaction effect for verb tense errors, which was the second error

type mentioned in the specific instructions. Our results did not support Proposition 2B, that the combination of specific instructions and good faith should lead to higher levels of compliance and performance. The interaction term of these variables was not statistically significant at traditional levels across all error types.

4. Discussion

4.1. The multiple facets of specificity

Dovetailing with the complex literature review we surveyed at the beginning of this paper, our findings show that any attempt to portray specificity as either good or bad might derive from a partial view of specificity. In contrast to most of the current views of specificity which focus either on the informative value (e.g. Schlag 1985; Kaplow 1992; Sunstein 1995) or the motivational aspect of specificity (see, e.g. Chou *et al.* 2010 for a behavioral perspective, or Shiffrin (2010) for a philosophical one), our design has attempted to explore both the motivational and the informative functions of specificity, using one integrative design featuring a few incentive compatible behavioral measures. Furthermore, we have taken a mild, realistic, and balanced approach, where we did not focus on the extreme high and low values of specificity but rather on the realistic hybrid version, in which vague legal standards are reinforced with some detailed examples. From our findings, we can begin to address what is the right approach to account for the different functions of legal specificity.

First, we found that respondents who were given specific instructions related to the editing task were more likely to attempt to correct errors found in the document than those who were given ambiguous instructions. Furthermore, we found that subjects who were given the specific instructions attempted significantly more corrections of error types that were mentioned in the instructions, as well as of error types that were not mentioned.

Second, providing detailed examples might be needed both for informativeness and for motivation. Our findings seem to suggest that when specificity provides useful information to participants, the motivational role of specificity is more indirect. We found that those who were given specific instructions were much more likely to attempt corrections for many error types relative to those who were given ambiguous instructions.

In addition to the ability of specificity to contribute to compliance, our findings show that it even led people to engage in better performance in areas beyond what was requested. By our interpretation, specificity can serve an informative function not only in the details it provides but also in its ability to help people focus on better performance, even in other categories of behavior.

When it comes to the ability of specificity design to help reduce the negative motivational effects, which were shown in earlier studies on specificity, a few conclusions can be reached from our study.

First, in contrast to the leading work of Chou *et al.* (2010), who argued that specificity increased resentment and consequently decreased performance, our study (which has focused on different circumstances) has shown that when specificity is phrased in an informative way, those resentful feelings are not triggered. Nevertheless, it still might be the case that when legal instructions are framed using highly constraining language, then the effect of specificity is indeed negative. Second, it is possible to employ specificity with examples rather than a closed list and thereby avoid many of the problems associated with over-focusing and lack of flexibility. Third, if one is interested in optimizing specificity in contracts as well as in law and regulation, multiple factors need to be taken into account, such as the relative costs of over-performance and under-performance, the consequences of mistakes, the cost of monitoring, and the information both the principal and the agent are expected to have about the task. Most of the answers to such questions require more focused empirical investigation and normative discussion, which need to take place with regard to the specific doctrine in question.

4.2. Good faith and specificity

Another important legal implication is related to the positive effect of good faith on people's interpretation of what is required from them. From our results, it seems that good faith was able to curb some of the failures associated with both specificity and monitoring. In particular, those exposed to the good faith treatment were less susceptible to the negative effect of monitoring. Given the previous results in the literature, our result suggests that an admonition for good faith may dispel distrust and resentment.

Naturally, there are other factors that should be taken into account, such as the need for information on the part of those receiving directions, and whether the examples provide information by giving the interpreter the correct signals regarding what would count as other examples. The examples we used were from the core definitions of editing, but participants still interpreted the request as a requirement to include even more remote examples of editing, including content-related mistakes that were planted in the text.

4.3. Is there an optimal balance of specificity, good faith, and monitoring?

In keeping with the exploratory nature of this study, the conclusions we can draw for legal design are open-ended. Because law covers a wide variety of contexts – in which specificity will sometimes be informative and sometimes, in keeping with other literature, may signal distrust – attention to which context is which is a pressing necessity (see also Feldman & Smith 2014). Those of our findings that most closely examine the hybrid catalog approach to regulation suggest the advantage of this approach where participants were not constrained by the examples provided and also focused on fixing errors that were not mentioned. Hence, in that regard, our findings corroborate the advantages of the catalog approach argued for by Parchomovsky and Stein (2015) as balancing the pros and cons of specificity.

The area that is probably closest to the editing task in our studies is contracts (although in our context there was no consent, and the source of the legal directive resembles an employment setting more than a general rule governing a situation, e.g. driving a car). The doctrine and the theoretical literature have long wavered on the benefits and costs of specificity and vagueness, both in contractual terms and off-the-rack rules (and standards) (see, e.g. [Schwartz & Scott 1995](#); [Choi & Triantis 2008, 2010](#)). More recently, experiments are being conducted to test the effect of specific goals for performance (see, e.g. [Brooks *et al.* 2014](#)). At the same time, good faith is a much-discussed, non-waivable feature that law reads into contracts. The main criticism of good faith is that it lacks content (see, e.g. [Miller & Perry 2012](#)). The formulation employed in our studies keys off the contractual notion of good faith, and the effects we found suggest that, however difficult it may be to implement good faith as a legal rule, it is quite likely to have important effects on contractual performance, which, after all, is the point of contracts.

4.4. Implications for other research paradigms

Because regulation varies in terms of specificity, monitoring, and good faith in ways that partially mirror the treatments in our study, our findings have implications for the literature on regulatory design. Existing taxonomies in regulation focus on the contrast between prescriptive and goal-based (and sometimes system-based) regulation, (e.g. [Coglianese & Lazer 2003](#); [Lodge 2004](#); [May 2007](#)). Prescriptive regulation sets forth – with some specificity – the details of what the regulated entity should do. And in goal-based regulation, the focus is not on what or how something is being done but rather on what is achieved by the regulated entity, which again requires choosing a level of specificity. In both kinds of regulation (and alternatives, like systems-based regulation), the level of specificity of directives is chosen against a backdrop of monitoring considerations and reliance on the good faith of the regulated actors.

We believe that the experimental and theoretical paradigm we offer to account for the level of specificity and how it interacts with good faith and monitoring contributes to this line of literature in a few ways. First, this framework highlights the behavioral components of the different types of instructions given to people. It suggests that when choosing the best regulatory approach, based on context, we also need to pay attention to the full range of people's likely reactions, such as crowding out and focus effects, an approach missing from the current regulatory literature. Whether specificity in any form of regulation supplies valued information or displaces intrinsic motivation might vary depending on whether the provided information relates to prescribed actions, required systems, or expected performance. Second, our findings suggest that even if one is interested in goal-based regulation, considerations of specificity might point, in some cases, to prescriptive regulation as the better tool to achieve the needed goal. Thus, based on an analysis of how much information the regulated actor has, the level of enforcement, and the actor's level of intrinsic motivation, our study might help to determine when regulators with an interest in achieving certain performance goals might adopt certain components of prescriptive regulation. Our result suggests that a focus in the regulatory literature on the interaction between specificity and the mode of regulation (prescriptive, system-based, performance-based) is likely to prove fruitful.

Because regulation shares features with other legal institutions, we believe that our framework has implications for these other aspects of law as well. Notably, tort law faces a choice between specifying conduct and holding actors responsible for results. For example, negligence specifies a standard of conduct (set out partly *ex ante* and partly *ex post*), while various forms of strict liability are framed not in terms of wrongful conduct but harmful results (see, e.g. Shavell 1980). (Admittedly, strict liability tends to be treated as more performance-based in economic models than in actual doctrine.) Within tort law, doctrines such as *res ipsa loquitur* (which raises a presumption of negligence based on results under circumstances that “speak for themselves”) and ultra hazardous activities (which holds actors strictly liable for harm resulting from activities like blasting) would be the closest to being performance-based. And both doctrines clearly raise issues of the interaction of specificity of the legal directive and monitoring of performance. One might even see a role for “good faith” in tort law, in its reliance on everyday moral standards. Thus, our results point toward design issues in those parts of the law that resemble the regulatory and contractual contexts.

To make the problem of specificity, detailed regulation, good faith, and monitoring more tractable, further work will need to be conducted in order to find out exactly which parts of the context are most important. Theoretical and empirical work may also suggest ways in which the law may address different audiences at different times: directives couched in terms of morality and good faith might be heard differently by those inclined to act ethically and those looking for loopholes to exploit (Feldman & Smith 2014). For now, one lesson to draw is that we cannot assume that specificity always crowds out intrinsic motivation or always ensures compliance – or always furnishes useful information to its addressees.

4.5. Limitations

While suggestive, especially against the background of the previous literature, our study is subject to limitations that caution against overgeneralization.

Theoretically, as the introduction to the paper has hinted, there is a wealth of theories that bear relevance to the likely effect of specificity on behavior. Some of the theories, mainly from rational choice and management literature on goal setting, suggest that ambiguous instructions might cause people to underperform. Some of the other theories that focus on crowding out motivation and trust see specificity as a device that will cause people to underperform because of the harm to the sense of autonomy and intrinsic motivation (e.g. Shiffrin 2010). The real challenge for legal policymakers would be to identify *ex ante* the contexts in which specificity will increase or decrease performance.

In addition to the general theoretical point, there are some methodological points that we need to acknowledge. Somewhat surprisingly and in contrast to previous studies, we found a lack of a monitoring effect. We identify two main methodological reasons, which might limit our ability to generalize from this lack of effect. First, given that our subjects were MTurk participants, even in the non-monitoring condition participants were already in a reputation market, especially in a task that allows the experimenter to potentially harm their MTurk reputation if they fail to recognize the errors in the document. Second, institutional review board limitations with regard to payments to participants on MTurk prevented us from more seriously threatening the payment to participants for failing to perform well. Focusing on the bonus payment allowed us some flexibility in “threatening” participants but this was a relatively limited threat and possibly not a credible one.

Another methodological limitation is related to the nature of copyediting. Originally we viewed this as a relatively straightforward task, in which people might have a limited need to get actual examples and directions as to what might be considered as editing, especially at the core level. Our results suggest that participants have come to use these examples on the informational level to a greater extent than we anticipated. Furthermore, it is possible that editing meant a lot to some people and less to others, which might have moderated the effect of specificity. Whereas for some participants, the specific instructions were helpful and informative and hence did not crowd out motivation, for others they were redundant and hence seemed to signal mistrust that might have created some motivational crowding out effect. While we have controlled for various factors that capture level of knowledge of English, no such control can be perfect. Future studies should also attempt to address the implications of the range and type of specificity we did – and did not – test. Based on research conducted in other contexts, it is possible to predict that across the range of specificities of law, we might be able to show a U-shaped effect, where with overly specific information we might see a backlash effect that might reduce performance. We focused on a hybrid approach to specificity (general standards with specific examples), which opens up the further possibility that different mixtures of types of information channels would lead to different behavioral effects. Future research might need to examine the possible effect of specificity in areas where participants

already know about the task and the effect of specificity would be mainly to reduce people's discretion and to improve the ability of the principal to monitor the behavior of the agent. Having said that, we believe that our findings do help to illuminate what we define as the multiple functions of legal specificity in regulating behavior.

In addition, it is not unwarranted to emphasize again the external validity limitations of our experimental design, especially with regard to state-driven regulations. We believe that our experiment, in which mTurk employees receive instructions of what is expected of them, how they will be monitored, and how much they will earn, is a realistic enough replication of important aspects of many contractual contexts. Because contract design is an important part of regulatory theory, we think that our contribution is significant even if it is limited to the intersection of private regulation and contract. The difference between the various modes of providing instructions is meant to correspond to the differences between possible contract designs. The editing task is similar enough, in our view, to the types of tasks one could contract over that they can provide a basis for testing the different modes of direction-giving. We do believe that additional studies should be conducted to replicate additional types of behavior and provide researchers with a broader picture of the effects of different types of regulatory design on people's behavior.

5. Conclusion

Specificity in directives of the kind employed by the law affects compliance and performance in context-sensitive ways, and our study is a first step in sorting them out. Contrary to arguments that specific instructions always crowd out motivation and harm performance, in our editing study, we find that, when guidance may be needed, adding specific examples to a vague standard may lead to better performance, even in contexts that extend beyond those examples. While the focus of this study was limited to a certain kind of specificity – a general rule with a list of examples – this experimental approach could be extended to other types of specificity and in other contexts of legal relevance. Overall, our study suggests that, although quite difficult, it is worthwhile for legal designers to think carefully about how to calibrate the cognitive and motivational effects of specificity. It seems that when the added information is provided in an informative way and especially with a focus on intrinsic motivation, the expected crowding out effect is counterweighted by the additional clarity added by providing specific examples as to how one should behave. Future research should move forward in accounting for the context that might tip the balance in different directions.

Acknowledgments

For helpful comments, the authors are grateful to Tami Kricheli-Katz, Maya Bar-Hillel, and participants at the Eighth Annual Conference on Empirical Legal Studies at the University of Pennsylvania, the 24th Annual Meeting of the American Law and Economics Association at the University of Chicago, the Second Meeting of the Private Law Consortium at the University of Pennsylvania, the 2013 American Law and Society Association Meeting, the behavioral law faculty workshop of the IDC, HUJI, and TAU. For excellent research assistance, we would like to thank Troy Schuller, Einav Tamir, Justin Zelin and Ryan Galisewski. We thank the Edmond J. Safra Center for the Study of Ethics at Harvard University and the support of the Israeli Science Foundation (grant 1283/11) for funding.

Notes

- 1 Please note that we also examine the effect of specificity on a different context, in which people need to evaluate their own behavior, as representing the situation of a sale of a product. The results of this experiment will be reported in a different study that focuses on the interaction between legal specificity and ethical decisionmaking.
- 2 Parchomovsky and Stein (2015, p. 191) offer the following example of a catalog regarding the meaning of prohibited dangerous weapons: “any pistol or other firearm (or imitation thereof) or other dangerous or deadly weapon (including a sawed-off shotgun, shotgun, machine-gun, rifle, dirk, bowie knife, butcher knife, switchblade knife, razor, blackjack, billy, or metallic or other false knuckles)” (quoting D.C. Code § 22-4502(a) (Supp. 2014)).
- 3 In all of the following propositions when we speak about “specificity” we adopt the perspective of specificity we refer to in the introduction – that is, adding a specific example to a vague legal ordinance.
- 4 One caveat is that, in many cases, external rewards can enhance intrinsic motivation. The interpersonal context in which the extrinsic motivation is introduced, or even the verbal cues attached to the sanctions, can determine how much we intrinsically value the extrinsic reward. For example, a child being reprimanded by a parent, whose opinion the child greatly values, may experience greater motivation to behave well than if the same reprimand were issued by a teacher with whom the child has little rapport. Nevertheless, the consensus in the literature suggests that in most instances, attempts to externally control people's behavior can have considerable counterproductive results in the long term (for a review of some of these conflicting effects, see [Deci *et al.* 2001](#)).

- 5 The results of an a priori power analysis (see Cohen 1988) suggest that a sample size of 339 and eight groups, held at an alpha level of 0.05, produces a power level $\pi = 0.451$ for a low effect size ($f = 0.1$) and $\pi = 0.996$ for a medium effect size ($f = 0.25$). Estimates of statistical power for different effect and sample sizes are illustrated in Appendix 6.4. As illustrated in the figure, while the power level ($\pi = 0.451$) for a low effect size ($f = 0.1$) given our sample size ($n = 339$) is indeed below traditional levels ($\pi = 0.8$), the power level for an effect size of $f = 0.15$ is essentially at traditional levels ($\pi = 0.786$). While not ideal, we are fairly confident that our sample size is able to detect smaller effect sizes.
- 6 Given that our study relies on a convenience sample, a discussion of the validity and reliability of our data is in order. There have been a number of recent studies, from a range of disciplines, which have evaluated the appropriateness of MTurk samples for experimental research (Buhrmester et al. 2011; Horton et al. 2011; Berinsky et al. 2012). With respect to external validity, there is reason to believe that MTurk samples indeed differ from nationally representative samples on some demographic characteristics (such as age and political ideology). However, using an MTurk sample, Berinsky et al. (2012) were able to replicate multiple experimental studies that have relied on both nationally representative and convenience samples. The authors conclude that, “if we treat the MTurk as a means for conducting internally valid experiments, instead of a representative sample, the MTurk respondent pool is very attractive” (Berinsky et al. 2012, p. 361). Thus, for the purposes of the present study, we are confident that the use of an MTurk sample is appropriate.
- 7 To demonstrate the validity of our measurement of task performance, namely, the number of attempted corrections made by the respondent, we calculated the correlation between time spent on the editing task and the percentage of total errors that were attempted to be corrected. We found a correlation coefficient of 0.43, indicating a moderately strong positive correlation between time and performance, suggesting that the focus on attempted corrections was the right one.

References

- Aquino K, Freeman D, Reed A II, Felps W, Lim VKG. (2009) Testing a Social-cognitive Model of Moral Behavior: The Interactive Influence of Situations and Moral Identity Centrality. *Journal of Personality and Social Psychology* 97, 123–141.
- Ayal S, Gino F. (2011) Honest Rationales for Dishonest Behavior. In: Mikulincer M, Shaver PR (eds) *The Social Psychology of Morality: Exploring the Causes of Good and Evil*, pp. 146–166. American Psychological Association, Washington, DC.
- Bardach E. (1989) Moral Suasion and Taxpayer Compliance. *Law & Policy* 11, 49–69.
- Berinsky AJ, Huber GA, Lenz GS. (2012) Evaluating Online Labor Markets for Experimental Research: Amazon.com’s Mechanical Turk. *Political Analysis* 20, 351–368.
- Bowles S. (2008) Policies Designed for Self-Interested Citizens May Undermine ‘The Moral Sentiments’: Evidence from Economic Experiments. *Science* 320, 1605–1609.
- Brooks RRW, Stremitzer A, Tontrup SW (2014) Stretch It But Don’t Break It: The Hidden Risk of Contract Framing. UCLA School of Law, Law-Econ Research Paper No. 13–22 (Last accessed 13 Nov 2013). Available from URL: <https://ssrn.com/abstract=2353733>
- Buhrmester M, Kwang T, Gosling SD. (2011) Amazon’s Mechanical Turk: A New Source of Inexpensive, Yet High-quality, Data? *Perspectives on Psychological Science* 6, 3–5.
- Choi A, Triantis G. (2008) Completing Contracts in the Shadow of Costly Verification. *The Journal of Legal Studies* 37, 503–534.
- Choi A, Triantis G. (2010) Strategic Vagueness in Contract Design: The Case of Corporate Acquisitions. *Yale Law Journal* 119, 848–924.
- Chou EY, Halevy N, Murnighan JK (2010) The Hidden Cost of Contracts on Relationships and Performance. IACM 23rd Annual Conference Paper. [Last accessed 27 Nov 2016.] Available from URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1612376
- Coglianese C, Lazer D. (2003) Management-based Regulation: Prescribing Private Management to Achieve Public Goals. *Law & Society Review* 37, 691–730.
- Coglianese C, Nash J, Olmstead T. (2003) Performance-based Regulation: Prospects and Limitations in Health, Safety, and Environmental Regulation. *Administrative Law Review* 55, 705–729.
- Cohen J. (1988) *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. Lawrence Earlbaum Associates, Hillsdale, NJ.
- Coletti AL, Sedatole KL, Towry KL. (2005) The Effect of Control Systems on Trust and Cooperation in Collaborative Environments. *Accounting Review* 80, 477–500.
- Cooter R. (1984) Prices and Sanctions. *Columbia Law Review* 84, 1523–1560.
- Cooter R. (2000) “Do good laws make good citizens? An economic analysis of internalized norms”. *Virginia Law Review* 1577–1601.
- Craswell R, Calfee JE. (1986) Deterrence and Uncertain Legal Standards. *Journal of Law, Economics, & Organization* 2, 279–313.
- Dana J, Weber RA, Kuang JX. (2007) Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness. *Economic Theory* 33, 67–80.
- Deci EL. (1971) Effects of Externally Mediated Rewards on Intrinsic Motivation. *Journal of Personality and Social Psychology* 18, 105–115.
- Deci EL, Koestner R, Ryan RM. (1999) A Meta-analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. *Psychological Bulletin* 125, 627–668.

- Deci EL, Koestner R, Ryan RM. (2001) Extrinsic Rewards and Intrinsic Motivation in Education: Reconsidered Once Again. *Review of Educational Research* 71, 1–27.
- Dickinson D, Villeval MC. (2008) Does Monitoring Decrease Work Effort?: The Complementarity Between Agency and Crowding-out Theories. *Games and Economic Behavior* 63, 56–76.
- Erev I, Ingram P, Raz O, Shany D. (2010) Continuous Punishment and the Potential of Gentle Rule Enforcement. *Behavioural Processes* 84, 366–371.
- Falk A, Kosfeld M. (2006) The Hidden Costs of Control. *American Economic Review* 96, 1611–1630.
- Fehr E, Falk A. (2002) Psychological Foundations of Incentives. *European Economic Review* 46, 687–724.
- Fehr E, Gächter S. (2000) Fairness and Retaliation: The Economics of Reciprocity. *Journal of Economic Perspectives* 14, 159–181.
- Feldman Y. (2011) “The complexity of disentangling intrinsic and extrinsic compliance motivations: Theoretical and empirical insights from the behavioral analysis of law”.
- Feldman Y. (2014) Behavioral Ethics Meets Behavioral Law and Economics. In: Zamir E, Teichman D (eds) *The Oxford Handbook of Behavioral Economics and the Law*, pp. 213–240. OUP, New York.
- Feldman Y, Smith HE. (2014) Behavioral Equity. *Journal of Institutional and Theoretical Economics* 170, 137–159.
- Feldman Y, Tyler TR. (2012) Mandated Justice: The Potential Promise and Possible Pitfalls of Mandating Procedural Justice in the Workplace. *Regulation & Governance* 6, 46–65.
- Ferguson MF, Peters SR (2000) But I Know it When I See it: An Economic Analysis of Vague Rules. [Last accessed 27 Nov 2016.] Available from URL: <http://ssrn.com/abstract=218968>
- Frank RH. (1988) *Passions Within Reason: The Strategic Role of the Emotions*. Norton, New York.
- Garcia SM, Chen P, Gordon MT. (2014) The Letter versus the Spirit of the Law: A Lay Perspective on Culpability. *Judgment and Decision Making* 9, 479–490.
- Garoupa N. (1997) The Theory of Optimal Law Enforcement. *Journal of Economic Surveys* 11, 267–295.
- Garoupa N. (2003) Behavioral Economic Analysis of Crime: A Critical Review. *European Journal of Law and Economics* 15, 5–15.
- Gibbons R, Henderson R. (2013) What Do Managers Do? Exploring Persistent Performance Differences Among Seemingly Similar Enterprises. In: Gibbons R, Roberts J (eds) *Handbook of Organizational Economics*, pp. 680–731. Princeton University Press, Princeton, NJ.
- Haisley EC, Weber RA. (2010) Self-serving Interpretations of Ambiguity in Other-regarding Behavior. *Games and Economic Behavior* 68, 614–625.
- Holmstrom B, Milgrom P. (1991) Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, & Organization* 7 (Special Issue, January), 24–52.
- Horton JJ, Rand DG, Zeckhauser RJ. (2011) The Online Laboratory: Conducting Experiments In A Real Labor Market. *Experimental Economics* 14, 399–425.
- Hsee CK. (1995) Elastic Justification: How Tempting but Task-irrelevant Factors Influence Decisions. *Organizational Behavior and Human Decision Processes* 62, 330–337.
- Kaplow L. (1992) Rules versus Standards: An Economic Analysis. *Duke Law Journal* 42, 557–629.
- Kim WC, Mauborgne R. (1996) Procedural Justice and Managers’ In-Role and Extra-Role Behavior: The Case of the Multinational. *Management Science* 42, 499–515.
- Kirchler E, Hoelzl E, Wahl I. (2008) Enforced versus Voluntary Tax Compliance: The “Slippery Slope” Framework. *Journal of Economic Psychology* 29, 210–225.
- Latham GP, Yukl GA. (1975) Assigned versus Participative Goal Setting with Educated and Uneducated Woods Workers. *Journal of Applied Psychology* 60, 299–302.
- Lazear EP. (2000) The Power of Incentives. *American Economic Review* 90, 410–414.
- Locke EA, Latham GP, Erez M. (1988) The determinants of goal commitment. *Academy of Management Review* 13.1, 23–39.
- Lodge M. (2004) Accountability and Transparency in Regulation: Critiques, Doctrines and Instruments. In: Jordana J, Levi-Faur D (eds) *Politics of Regulation*, pp. 124–144. Edward Elgar Publishing, Cheltenham, UK.
- Logue KD. (2007) Optimal Tax Compliance and Penalties When the Law Is Uncertain. *Virginia Tax Review* 27, 241–296.
- Marshall M, Harrison S. (2005) It’s About More Than Money: Financial Incentives And Internal Motivation. *Quality & Safety in Health Care* 14, 4–5.
- May PJ. (2007) Regulatory Regimes and Accountability. *Regulation & Governance* 1, 8–26.
- Mazar N, Amir O, Ariely D. (2008) The Dishonesty of Honest People: A Theory of Self-concept Maintenance. *Journal of Marketing Research* 45, 633–644.
- Miller AD, Perry R. (2012) Good Faith Performance. *Iowa Law Review* 98, 689–745.
- Osterloh M, Frey B. (2013) Motivation Governance. In: Grandori A (ed) *Handbook of Economic Organization: Integrating Economic and Organization Theory*, pp. 26–40. Edward Elgar, Cheltenham, UK.

- Parchomovsky G, Stein A. (2015) Catalogs. *Columbia Law Review* 115, 165–209.
- Poppo L, Zenger TR. (2002) Do Formal Contracts and Relational Governance Function as Substitutes or Complements? *Strategic Management Journal* 23, 707–725.
- Prendergast C. (1999) The Provision of Incentives in Firms. *Journal of Economic Literature* 37, 7–63.
- Schlag PJ. (1985) Rules and Standards. *UCLA Law Review* 33, 379–430.
- Schwartz A, Scott RE. (1995) Political Economy of Private Legislatures. *University of Pennsylvania Law Review* 143, 595–654.
- Seijts GH, Latham GP. (2001) The Effect Of Distal Learning, Outcome, And Proximal Goals On A Moderately Complex Task. *Journal of Organizational Behavior* 22, 291–307.
- Shalvi S, Gino F, Barkan R, Ayal S. (2015) Self-Serving Justifications: Doing Wrong and Feeling Moral. *Current Directions in Psychological Science* 24, 125–130.
- Shavell SM. (1980) Strict Liability versus Negligence. *The Journal of Legal Studies* 9, 1–25.
- Shavell SM. (2002) Law versus Morality as Regulators of Conduct. *American Law and Economics Review* 4, 227–257.
- Shiffrin SV. (2010) Inducing Moral Deliberation: On the Occasional Virtues of Fog. *Harvard Law Review* 123, 1214–1247.
- Sunstein CR. (1995) Problems with Rules. *California Law Review* 83, 953–1026.
- Tenbrunsel AE, Messick DM. (1999) Sanctioning Systems, Decision Frames, and Cooperation. *Administrative Science Quarterly* 44, 684–707.
- Tyler TR. (2006) *Why People Obey the Law*. Princeton University Press, Princeton, NJ.
- Tyler TR, Darley JM. (2000) Building a Law-Abiding Society: Taking Public Views about Morality and the Legitimacy of Legal Authorities into Account When Formulating Substantive Law. *Hofstra Law Review* 28, 707–739.
- Welsh DT, Ordóñez LD. (2014) Conscience without Cognition: The Effects of Subconscious Priming on Ethical Behavior. *Academy of Management Journal* 57, 723–742.
- Wenzel M. (2005) Motivation or Rationalization? Causal Relations between Ethics, Norms and Tax Compliance. *Journal of Economic Psychology* 26, 491–508.
- Zhou J, Shalley CE. (2003) Research on employee creativity: A critical review and directions for future research. *Research in Personnel and Human Resources Management* 22, 165–218.

Appendix

A.1. Summary Statistics

Specificity	Treatment		Narrow Compliance		Broad Compliance					Performance	
	Good faith	Verification	Verb Tense	Punctuation	Subject-Verb Agreement	Missing Words	Word Meaning	Pronouns	Articles	Singular/Plural	Apostrophes
1	1	1	5.171 (3.451) 35	4.600 (2.569) 35	6.943 (3.662) 35	2.771 (2.102) 35	2.629 (1.880) 35	0.600 (0.695) 35	0.314 (0.471) 35	0.257 (0.443) 35	0.514 (0.507) 35
1	1	0	5.042 (3.377) 48	5.229 (2.146) 48	7.542 (3.235) 48	3.229 (2.205) 48	2.729 (1.710) 48	0.646 (0.601) 48	0.354 (0.483) 48	0.313 (0.468) 48	0.625 (0.489) 48
1	0	1	4.923 (3.301) 52	4.808 (2.179) 52	7.096 (3.133) 52	2.365 (2.133) 52	2.250 (1.667) 52	0.481 (0.505) 52	0.308 (0.466) 52	0.288 (0.457) 52	0.365 (0.486) 52
1	0	0	5.163 (3.147) 43	5.442 (2.260) 43	7.395 (3.156) 43	3.302 (2.099) 43	2.837 (1.463) 43	0.674 (0.606) 43	0.279 (0.454) 43	0.302 (0.465) 43	0.581 (0.499) 43
0	1	1	3.952 (3.414) 42	4.643 (2.497) 42	5.714 (3.909) 42	2.333 (2.032) 42	2.286 (1.916) 42	0.690 (0.715) 42	0.405 (0.497) 42	0.238 (0.431) 42	0.405 (0.497) 42
0	1	0	3.463 (2.916) 41	4.293 (2.205) 41	5.805 (3.311) 41	2.146 (1.726) 41	1.976 (1.255) 41	0.415 (0.547) 41	0.244 (0.435) 41	0.171 (0.381) 41	0.390 (0.494) 41
0	0	1									

(Continues)

Treatment			Narrow Compliance		Broad Compliance					Performance	
Specificity	Good faith	Verification	Verb Tense	Punctuation	Subject-Verb Agreement	Missing Words	Word Meaning	Pronouns	Articles	Singular/Plural	Apostrophes
			3.435 (2.934)	4.674 (2.395)	5.043 (3.584)	2.022 (1.938)	1.913 (1.427)	0.522 (0.586)	0.348 (0.482)	0.130 (0.341)	0.457 (0.504)
0	0	0	46	46	46	46	46	46	46	46	46
			4.188 (2.901)	4.531 (2.170)	6.125 (3.635)	2.906 (1.748)	2.188 (1.615)	0.594 (0.499)	0.281 (0.457)	0.156 (0.369)	0.500 (0.508)
			32	32	32	32	32	32	32	32	32

Treatment conditions: 1 = Yes, 0 = No. Summary statistics include cell means, standard deviations in parentheses, and number of respondents in each cell.

A.2. Treatment prompt wording

Treatment	Prompt
Specificity	<p>We would like you to carefully edit a text document to improve its readability and quality. When editing the document, the typical types of errors to which you should pay particular attention include, but are not limited to, mistakes in punctuation and verb tense. (I understand and agree.)</p> <p>Your task is to edit a document which contains, among others, punctuation and verb tense sequencing errors. Below are some examples of proper punctuation and verb tense usage.</p> <p>Punctuation:</p> <p>Incorrect: While Jane loves chocolate, she bought vanilla ice cream today. Correct: While Jane loves chocolate; she bought vanilla ice cream today.</p> <p>Incorrect: How are you feeling today, Jake. Correct: How are you feeling today, Jake?</p> <p>Incorrect: He plays basketball. His passion however lies with soccer. Correct: He plays basketball. His passion, however, lies with soccer.</p> <p>Verb Tense:</p> <p>Incorrect: After he had finished his meal, John <i>would pay</i> the bill. Correct: After he had finished his meal, John <i>paid</i> the bill.</p> <p>Incorrect: Most current car designs <i>are providing</i> air conditioning. Correct: Most current car designs <i>provide</i> air conditioning.</p> <p>Incorrect: Linguists agree that the Finnish language is closely <i>relating</i> to Estonian. Correct: Linguists agree that the Finnish language is closely <i>related</i> to Estonian.</p> <p>Please complete the following exercises on proper verb tense sequencing and punctuation usage before editing the document:</p> <ul style="list-style-type: none"> • Roger _____ to the supermarket the other day to buy some bread. (has gone/had gone /went/goes) • While he _____ in line at the cashier, his girlfriend called him on his mobile phone. (waited/was waiting/wait/had waited) • John told Jane exactly what she wanted to hear that she plays the violin well. (what she wanted to hear, that she/what she wanted to hear: that she/what she wanted to hear; that she/No punctuation correction needed.) • His favorite movies are <i>Casablanca The Godfather Pulp Fiction</i> and <i>Star Wars</i>. (are <i>Casablanca, The Godfather, Pulp Fiction</i> and <i>Star Wars</i>./are <i>Casablanca, The Godfather, Pulp Fiction</i> and <i>Star Wars</i>./are; <i>Casablanca, The Godfather, Pulp Fiction</i> and <i>Star Wars</i>./No punctuation correction needed.) <p>Are spelling mistakes one of the errors that should be corrected? (yes/no)</p>
Ambiguous	<p>We would like you to carefully edit a text document. We are not giving you specific examples of what kind of errors you should focus on. After reading the document completely, please edit and correct any errors in order to improve the quality and readability of the document. (I understand and agree.)</p> <p>Should my edited work be clear and easily read? (yes/no)</p>
Good faith	<p>You are expected to act in <i>good faith</i> when completing the tasks. This means that you should act in a genuine, sincere, fair, and honest manner while participating in this study. (I understand and agree to act in good faith while completing the editing task.)</p> <p>Please define in your own words (1–2 sentences) what <i>good faith</i> means to you.</p> <p>Please define in your own words (1–2 sentences) what it means to edit a document in a good faith.</p>
No good faith	[No text]
Verification	<p>Please note that satisfactory execution of the editing instructions will be rewarded with an extra \$1.00 bonus payment. Prior to paying you the bonus, we will carefully review your performance as well as how closely you followed our instructions. Your bonus payment will</p>

(Continues)

Treatment	Prompt
	<p>be dependent on how well you followed our instructions. (I understand and agree that the quality of my editing will be assessed after submission.)</p> <p>How many minutes do you think the researchers will invest in verifying that you followed the instructions prior to deciding whether to pay you the \$1.00 bonus? (open-ended.)</p>
No verification	[No text]

Wording of all treatment prompts prior to editing task. Subjects were shown a combination of the treatment prompts presented in this table depending on which conditions they were assigned to. Required responses to the prompts are displayed in parentheses. Subjects who were not assigned to the good faith or verification treatment conditions were not shown additional text.

A.3. Document for Editing

What, then, are the general justifications for a legal institution that gives exclusive rights to persons over the things they own² A wide variety of justifications for private property have been offered. We will mention **five**².

First, the institution of property provides an effective way of managing **societies**⁹ resources. We can imagine a variety of techniques for managing resources, ranging from governments bureaucracies to local customs to a system of might-makes-right. The property strategy for resource management can be seen as one that delegates near-dictatorial powers over particular resources to individual owner-managers, which powers are then **backed up the**⁴ authority of the state. This strategy, to a greater degree than government **bureaucracy**² allows for decentralization in the management of resources, and **permit**³ the owner-managers to specialize in developing the knowledge and skills pertinent to their particular resource. The decentralized nature of property also permits experimentation with new uses and techniques for managing resources, rendering it more dynamic than a consensus-based system of local custom. And, **so long the**⁴ state provides a sufficient degree of security for ownership, the property strategy **are**³ more stable and requires fewer expenditures on defensive measures than does might-makes-right system.

Third⁵, the institution of property provides a powerful set of incentives for persons to make investments in and engage in effective management of the resources they control. A homey way of expressing this is that property allows the owners to reap where they have sown. Farmers who own land, for example, can decide what kind of seeds to plant, how much effort to put into cultivating them, and when and how to harvest the crops, knowing that the effort and skill they put into the process will be reflected in the yield they ultimately **obtain**² Other forms of resource management create much less incentive to expend these kinds of efforts to make productive use of resources, because they provide no assurance that those who put in effort to improve the resource will be able to appropriate the fruits of their efforts. This is because no one is in a position to exclude others from interfering with or seeking to capture the benefits of their efforts.

Third, the institution of property facilitates the making of contracts regarding the use and control of resources. In order to contract for the exchange of or modifications to the use of resources, it is necessary to know who controls what resources and, hence, who may make contracts with respect to those resources. Property gives us the answer to these questions. Two types of contracts regarding property **were**¹ especially important. One is the exchange of property rights. Suppose I own Blackacre but **is**³ too old and tired to make much use of it. If I can enter into a contract to sell it to someone else, then this resource may be used more effectively, making society better off. Another is a contract modifying the use of property. If my neighbor is using **their**⁶ property in a way that **cause**³ me discomfort or irritation, one solution is to enter into a contract with the neighbor in which he or she promises to desist from this use.

Fourth, property is an important source of individual autonomy. Property provides the material means for individuals to achieve a degree of independence from others. By giving individuals control over resources, it allows them to control the direction of their lives. Particular items of property may also be critical to personal identity or to the development of individual personality. Many peoples' identity is closely wrapped up with their homes. For others, particular shops or factories may be vital to how they see themselves. Most of us have books, photos, memorabilia, items of clothing, or collections of things to which we attach significance, and **where**⁵ we would be pained to lose.

Fourth⁵, property **are**³ important to the preservation of liberty. Morris Cohen once **written**¹ that property is a form of sovereignty; the right to exclude others from things is source of power over other people. This is true, but it also means that if the ownership of property is distributed widely enough, then property ownership can be a source of countervailing power to the power of the government—or the power of other property owners. Checks and balances

are thought to be a vital to preserving liberty, and dispersed property ownership **provide**³ an important source of checks and balances. For example private property allows individuals to organize opposition parties and distribute literature critical of the government; if the government controls all the resources it is easy to squelch dissent. And, property **allow**³ unpopular minorities to resist threats from the government to cut off government-granted **benefit**⁸ or employment.

Although property has been applauded for these positive functions, certain general concerns have also been raised by the institution of property.

One pervasive problem goes by the name “externalities.” The private property strategy entails dividing the world **down**⁵ into separate parcels of land and discrete **objects personal**⁴ property, each with its individual owner. But the owner-managers of these individualized units of property may **uses**¹ them in ways that have spillover effects for the owner-managers of other units of property. Spillover effects that **had**¹ adverse consequences for others, known as negative externalities, **are particular**⁴ matter of concern. A can use his land in a way that generates pollution damaging to his neighbor B. Or C can allow her car to deteriorate into an unsafe condition, posing a danger on the road to D and other drivers. In fact, the very strategy of allowing owners to appropriate the **losses**⁵ from their property creates an incentive to try to foist as many of the costs associated with property as possible onto someone else. Any set of legal rules governing property will thus **have come**⁴ up with strategies for trying to control negative externalities.

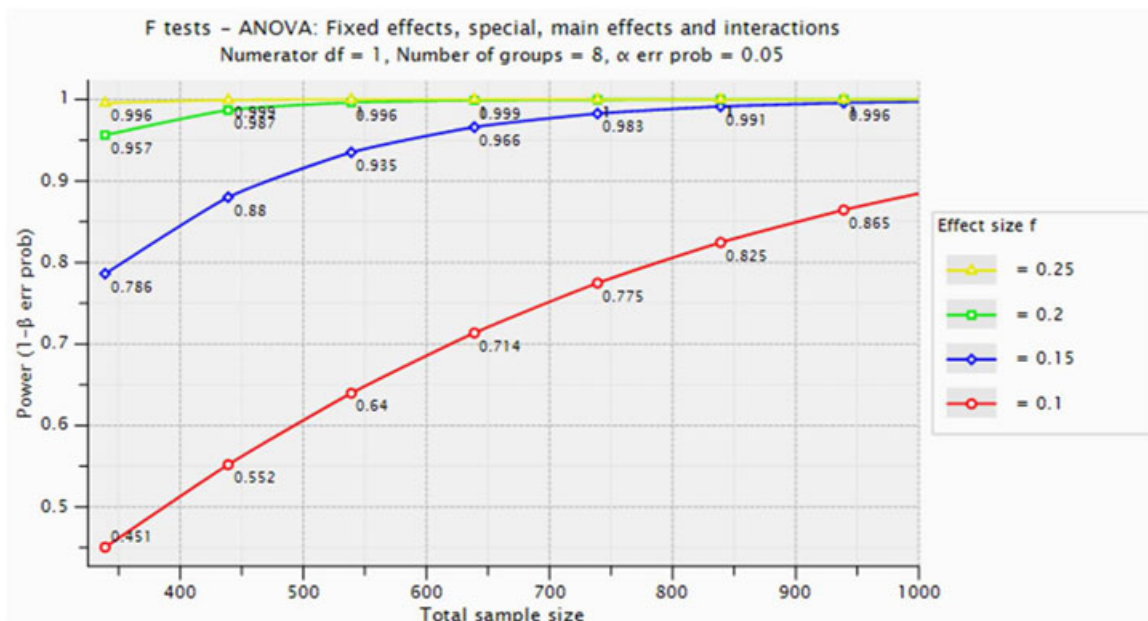
Another concern about property is monopoly. Property, by its very nature, confers a monopoly of control on someone with respect to a particular resource. Every property right is in this sense a monopoly right. Ordinarily, this is **of concern**⁴. For example each farmer has a monopoly on his or her own land. But if there are thousands of farmers producing a substantially identical commodity—**say**² wheat—there will be vigorous competition among the farmers in the market for wheat, and the monopoly each farmer has over his or her own production facilities will have no effect on the price that consumers must pay for wheat. In other circumstances, **however**² granting property rights can create monopolies that do have troublesome social consequences. For example, if I own the only piece of land on which it is feasible to build a bridge across a river, this may give me a monopoly on river-crossings, and with it the power to extract large **polls**⁵ from the public for the privilege of using the bridge. **Similarly**² the award of patent rights or copyrights may in certain circumstances allow the owners of these rights to extract very large payments from the public, if there are no good substitutes for the thing they have created. This concern therefore **suggested**¹ that the property strategy **will be becoming**¹ problematic insofar as the monopoly rights conferred by property **confess**⁵ with a distinct market for particular goods or services. It may be necessary in these circumstances to modify property rights by applying antitrust laws or some form of **regulation**².

Another concern about property is that it leads to commodification of values and social relations. Property conceives of the world in terms of owners dominating or controlling objects. Many people resist thinking this way about their bodies, their intimate relations, their networks of friends and **colleagues**² their pets, and so forth. The concern here is with the scope or domain of property rights. The more we extend the sphere of property, the more we **thought**¹ about the world in terms of owner-object relations. Insofar as it is important to preserve a sphere of life that is organized according to different principles, then we must **have exercised**¹ caution about how far the system of property rights extends.

Finally, property **have**³ long been attacked on the ground that it **promoting**¹ inequality. Property, at least if it is well managed, **tend**³ to beget more property. This is because property, by allowing the **land**⁵ to exclude others, **permitted**¹ the owner to capture the fruits of the property, to reap what **have**³ been sowed. Much of the captured “fruit” is attributable the skill and industry of the owner—but not all. Some is attributable to rising demand for resources generally, and to sheer luck. The component that can be ascribed to luck or general conditions of scarcity represents a kind of built-in multiplier whereby those who have property get more property, without regard to **them**⁶ individual desert. If we combine this with a general right of inheritance, then a robust system of private property can create conditions where the rich generally get richer. This is not to say that abolishing property would create greater equality. Communist systems **was**³ notorious for providing special perks for party leaders that gave **them lavish**⁴ lifestyle **for**⁵ beyond what ordinary workers could aspire enjoy. But, it does suggest a rationale for imposing higher burdens of taxation on those with significant property to offset the dynamic tendency toward ever greater **inequality**² Whether this greater burden should be greater than proportionate with wealth or income (i.e. progressive), **and so**⁴, how much so, **has**³ been a topic of lively debate—as have the question of whether to use taxes or the rules of property itself to achieve distributive ends.

Superscripts refer to the following error type categories: 1 = Verb tense (10); 2 = Punctuation (10); 3 = Subject-Verb Agreement (12); 4 = Missing words (8); 5 = Meaning of words/sentences (9); 6 = Pronoun Agreement Error (2); 7 = Incorrect article usage (1); 8 = Singular/Plural (1); 9 = Improper Apostrophe Usage (1)

A.4. Estimates of statistical power



The results of an a priori power analysis with eight groups and an alpha level of 0.05 are illustrated in this figure. Power estimates are calculated for effect sizes ranging from “low” ($f = 0.1$) to “medium” ($f = 0.25$) at intervals of 0.05 (see Cohen 1988). ANOVA, analysis of variance.