THE QUALITY OF GOVERNMENT INSTITUTE

UNIVERSITY OF GOTHENBURG

Legal Clarity and Impartiality:

A Global Experimental Study of Consistency in Bureaucratic Decision-Making

2025

Working paper series 2025:5

Joakim Nilsson
Marina Nistotskaya

THE QUALITY OF
GOVERNMENT INSTITUTE

GÖTEBORGS
UNIVERSITET

Legal Clarity and Impartiality:

A Global Experimental Study of Consistency in Bureaucratic Decision-Making

Joakim Nilsson

Marina Nistotskaya

WORKING PAPER SERIES 2025:5

QoG THE QUALITY OF GOVERNMENT INSTITUTE
Department of Political Science
University of Gothenburg
Box 711, SE 405 30 GÖTEBORG
August 2025
ISSN 1653-8919

Legal Clarity and Impartiality: A Global Experimental Study of Consistency in Bureaucratic Decision-Making
Joakim Nilsson
Marina Nistotskaya
QoG Working Paper Series 2025:5
August 2025
ISSN 1653-8919

Abstract

The language used in legal texts is often ambiguous, hindering bureaucrats' ability to understand, interpret, and apply the law consistently, and thereby threatening impartiality. While the Quality of Government (QoG) literature emphasizes the importance of impartiality, it overlooks how the clarity of legal language shapes this principle in practice. This paper bridges two bodies of scholarship: the QoG literature and legal studies that highlight the role of language clarity in law comprehension but haven't considered its implications for bureaucratic decision-making. We advance the argument that language clarity fosters impartiality by enabling more consistent application of the law. To test this claim, we conducted an online survey experiment with 900 current and former government officials from 33 countries. Participants were randomly assigned to resolve a case – based on a real-life scenario – in which the legal provision was presented in either ambiguous or clearer language. The results show that exposure to ambiguous wording reduced consistently in the application of the law, whereas clear language fostered greater judgment consistency. These results call for a revision of the prevailing conceptualization of high QoG – from a mere absence of factors not "stipulated in the policy or the law" to also include the clarity with which laws *are* formulated. The paper underscores the practical significance of legal drafting for public sector performance.

Key words: language clarity, impartiality, bureaucratic decision-making, consistency, experiment.

We gratefully acknowledge the financial support of the Quality of Government Institute, University of Gothenburg.

Joakim Nilsson
joakimn.1993@gmail.com

Marina Nistotskaya
The Quality of Government Institute
Department of Political Science
University of Gothenburg
marina.nistotskaya@gu.se

1. Introduction

Legal language is often unclear, hindering bureaucrats from correctly understanding, interpreting, and consistently applying laws. This, in turn, threatens impartial application of laws – a foundational pillar of high-quality government (Rothstein & Teorell, 2008).

While a large literature has examined the link between impartiality with a number of positive government outputs and various positive societal outcomes (Ahlerup et al. 2016; Boräng et al. 2018; Dahlberg & Holmberg 2014; Kolvani & Nistotskaya 2025; Nistotskaya et al. 2015; Nistotskaya & Cingolani 2016), the causes of bureaucratic impartiality have remained under researched. Specifically, extant research largely neglects the role of law in shaping impartiality. While legal origin theory addresses the role of legal systems in government performance (La Porta et al. 1999; La Porta et al. 2008), it falls short in accounting for legal clarity. Similarly, legal research demonstrates that clarity is essential for understanding and interpreting laws, but it limits its focus to the judicial sphere (Charrow & Charrow 1979; Diamond & Levi 1995; Randall 2014) and offers no insight into how legal clarity affects bureaucratic decision-making. These two strands of literature – QoG and law scholarship – remain largely disconnected: there is currently no research examining the link between legal clarity and bureaucratic impartiality. The aim of this paper is to fill in this research gap by investigating the link both theoretically and empirically.

Building on the legal linguistics literature, we develop a theoretical argument that greater linguistic clarity in legal texts leads to more consistent application of law across similar cases, thereby enhancing impartiality and test this proposition through an online survey experiment involving nearly 900 current and former government officials worldwide.

The paper progresses as follows: section 2 reviews the relevant literature and section 3 discusses the theoretical framework that connects legal clarity to impartiality. Section 4 outlines the study design, methodology, and process, followed by the presentation of the results of the analysis in section 5, discussion and conclusion.

2. Literature Review

2.1. Quality of Government as Impartiality

Over the past several decades, the field of comparative politics has evolved from the focus on institutions regulating access to political power (or the input side of the political system) to incorporate the output side of the political system – concerning how political power is exercised (La Porta et al. 2008; Rothstein & Teorell 2008). Within this, research has moved away from broad and often ambiguous concepts such as "good governance" to more precisely delineated concepts, thereby enhancing the conceptual clarity and measurability (Nistotskaya 2020). For example, the concept of the quality of government (QoG) has been introduced to focus on impartiality in the implementation of law: "When implementing laws and policies, government officials shall not take into consideration anything about the citizen/case that is not beforehand stipulated in the policy or the law" (Rothstein & Teorell 2008, p. 170).

While Rothstein and Teorell's (2008) definition of impartiality is analytically rigorous, the concept has been subject to critique. Some scholars argued that this perspective overlooks the role of the content of the laws, advocating for a broader conceptualization of impartiality – one that takes into account the normative content of laws, which is presumed to influence their outcomes (Agnafors 2013; Olander 2021; Sparling 2018). Although moral and ethical considerations may be important, they belong to a different debate. In Rothstein and Teorell's (2008) framework, impartiality is deliberately restricted to the application of laws, emphasizing the equal and unbiased implementation of already established rules by public officials.

We argue, however, that their definition contains a critical shortcoming: it presumes that laws and policies exhibit a high degree of legal clarity and precision that precludes ambiguity and misinterpretation, and no misunderstandings or interpretive challenges – resulting from legal ambiguity – arise that could jeopardize the impartial application of the law. In other words, Rothstein and Teorell's (2008) assume away the problem of linguistic clarity in legal texts. Since the clarity of the law cannot be taken for granted, it is essential to consider how it affects the practice of impartiality in real-life settings.

While Rothstein and Teorell's (2008) definition of impartiality emphasizes the absence of influence from factors not "stipulated in the policy or the law" (p. 170), this paper argues that the

clarity of what is stipulated in the law also plays a critical role in enabling impartial implementation. Given the foundational principle of impartiality that "we must treat like cases alike" (Dworkin 1986, p. 165), it becomes evident that the ambiguity of wording of laws (hereafter *legal ambiguity*) can jeopardize the treatment of like cases alike or, in other words, undermine impartiality. We argue that linguistic clarity in legal texts is likely to affect the consistency of the application of the law: while legal ambiguity leads to less consistency – and less impartiality, higher linguistic clarity leads to more consistent application of law/higher impartiality.

2.2. Legal Research

2.2.1. Legal Origin Theory

A prominent strand of research – known as Legal Origin Theory (LOT) – argues that different legal traditions have a lasting effect on government performance and societal outcomes. LOT distinguishes between common and civil law traditions,[1] with the former being associated with stronger property rights, lighter regulation, and better performance than later system (Glaeser & Shleifer 2002; La Porta et al. 1999; La Porta et al. 2008; Posner [1973] 2007). The LOT literature also observes that legal clarity may vary across legal traditions, suggesting that the civil law tradition may exhibit lower linguistic clarity. One possible reason is that statutes in this tradition often begin with general statements of principles, addressing intricate subject matter across diverse circumstances and intended for multiple audiences (Thomas 1985). By contrast, common law statutes typically target a more specific audience – informing directly those affected by the law of their rights and obligations – and aim to spell out as much detail as possible. While this approach produces greater specificity, it also results in highly complex and often lengthy texts (Dale, 1977; Vanterpool 2007). Both traditions remain susceptible to legalese – the specialized jargon of the legal profession – but notable efforts to counteract this through plain language reforms have occurred primarily in common law countries (Blasie 2022; Kimble 1992; Schriver 2017; Thomas 1985; Vanterpool 2007).

[1] Unlike civil law, common law is formed not only by statutes adopted by legislatures, but also through the rulings of appellate judges (Gibbons 1999; La Porta et al. 1999; Posner [1973] 2007).

2.2.2. Legal Comprehension

Although empirical studies directly testing the connection between legal clarity and impartiality are lacking, some insights can be drawn from legal scholarship that has examined legal comprehension. There is a small number of studies that have explored how the clarity of legal texts influence individuals' ability to understand legal texts.

In one of the first studies of this kind, Charrow and Charrow (1979) examined instructions for trial juries, showing that unclear legal language makes legal texts less comprehensible to laypeople working in courts.

Diamond and Levi (1995) showed that misunderstanding of legal text can have serious consequences. In U.S. courts, jurors without formal legal training play a significant role in decisions involving the death penalty. In their study, jurors recruited from a Chicago court were presented with hypothetical cases and legal instructions written in complex language. To assess their understanding of the instructions, participants were asked to reproduce parts of the legal instructions. The results indicate that many jurors had difficulty understanding essential aspects, which, in extreme cases, lead to erroneously imposed death sentences.

Similar findings were obtained by Randall (2014) from several experimental settings. In one experiment, participants were exposed to a sample of real-life jury instructions and had to respond to true/false questions, assessing their comprehension. The results revealed that understanding declined when the instructions included linguistically complex features – such as passive voice and presupposed information – both of which are linked to the increased cognitive processing load. In a follow-up experiment, the instructions were rewritten in plain English to eliminate these features, resulting in considerably improved comprehension.

The legal comprehension literature provides a point of departure for understanding the relationship between the extent of clarity of the language of legal texts and their comprehension. However, it addresses neither how legal clarity affects individuals' ability to treat "like cases alike" (Dworkin 1986, p. 165) – a core aspect of impartiality – nor how it affects this ability in the context of public administration.

2.2.3. Research Gap

Quality of government is predominantly understood as the impartial implementation of laws and policies by the bureaucracy (Rothstein & Teorell 2008). This conception emphasizes that bureaucracy acts as a neutral entity, ensuring that the realization of the legal norm is not subverted by external influence. This conception, however, overlooks the role of the law. Existing critique emphasizes the content of laws as an important factor of impartiality (Agnafors 2013), yet it pays no attention to the wording of laws – linguistic clarity. Similarly, a recent literature on regulatory burden argues that the volume and complexity of legislation hinders effective policy implementation (Fernández-i-Marín et al. 2024a, 2024b, 2024c; Duvanova 2017) but doesn't consider linguistic clarity as an element of legislative complexity.

On the other hand, the existing legal research engages with the law in several distinct ways. While some link government performance to broad differences in legal systems, there is also empirical research on legal comprehension showing that unclear wording reduces understanding of the legal provisions and increases the risk of misapplication. However, none of these strands directly examine how legal clarity affects impartiality of decision-making, particularly in the context of public administration.

In sum, existing literature overlooks the question of whether linguistic clarity of legal texts (thereafter legal clarity) affects the bureaucracy's ability to implement laws impartially. This paper seeks to close this gap by examining the impact of legal clarity on the consistency of decision-making by bureaucrats.

3. Theory

Our theoretical argument draws on legal linguistics research and on Dworkin's concepts of law as integrity and equality before the law (1977, 1986), which together imply that legal decisions must be consistent across cases – "we must treat like cases alike" (Dworkin 1986, p. 165). The legal comprehension literature (discussed above) maintains that clear language conveys the substance of the law more effectively, thereby enhancing implementers' understanding of legal provisions. We further argue that such improved understanding increases the likelihood that the law is applied more consistently, and thus more impartially.

In this paper, we examine the clarity in legal language, focusing on specific linguistic features, that legal linguistics research has identified as hindering comprehension and complicating interpretation. These features form the basis of our framework for assessing legal clarity, which we regard as higher when such features are absent.

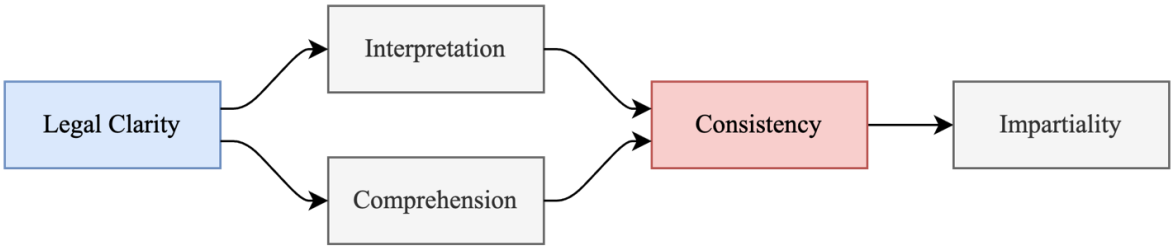
First, ambiguous pronoun placement can produce multiple interpretations, as in the sentence "John thought he should be more polite to Bill," where the referent of "he" is unclear (Solan 1995, p. 1070). Second, nominalization – the transformation of verbs into nouns – adds complexity to sentences by making them more abstract and dense (Charrow & Charrow 1979; Gibson 1998, Tiersma 1999; Williams 2004), as in "the determination was made" instead of the simpler "they determined". Nominization also obscures meaning because it removes the clear action and agent from the sentence, making it harder for readers to grasp who is doing what. Similarly, in passive voice constractions, the grammatical subject receives the action ("The contract was breached") rather than performing it, and the actor is frequently omitted altogether ("The contract was breached by X" → "The contract was breached"). This omission increases interpretive ambiguity and shifts cognitive effort to the reader, who must infer missing information (Charrow & Charrow 1979; Kimble 1994, 1998; Tiersma 1999). Passive voice also generally produces longer, wordier setences, slowing down processing speed and increasing the cognitive load, and increasing the risk of misinterpretation (Kimble 1994, 1998; Tiersma 1999; Williams 2004). Forth, embedded clauses, which stack multiple ideas within a single sentence, interrupt linear progression of thought, forcing readers to store several incomplete structures in working memory before they can resolve the sentence's meaning (Charrow & Charrow 1979; Gibson 1998; Tiersma 1999; Williams 2004). In legal language, this is especially problematic because embedded clauses often appear midway through complex noun phrases, such as "Any person, who having knowingly made a false statement in an application for a permit, shall be liable…". Psycholinguistic research shows that such center-embedded structures significantly increase cognitive load, lengthen reading time, and elevate error rates in comprehension, particularly when more than one level of embedding is used (Gibson 1998). Fifth, negative sentence structures, particularly those with multiple negatives, reduce clarity and increase cognitive load – for instance, "It is not uncommon not to know" is less clear than a straightforward positive construction (Charrow & Charrow 1979).

Additionally, certain terms are considered technical legal vocabulary to varying degrees. For example, while "credibility" is a common English word whose specialized legal meaning (related to the believability or trustworthiness of a witness or evidence) is generally accessible to non-jurists, "imputed" is a technical legal term – referring to attributing responsibility or knowledge to someone even if they did not directly perform the act or possess the knowledge – whose specialized meaning may be inaccessible to non-jurists. This variability in meaning may lead to misinterpretations by implementers (Charrow & Charrow 1979), contributing to inconsistency in the application of the law across cases. The same applies to Latin phrases, archaic or rarely used words and expressions that are frequently used in legal texts (Gibbons 1999, Williams 2004). Finally, *ambiguous formulations* hinder understanding. For example, "In our current circumstances, the budgetary aspect is a factor which must be taken into consideration to a greater degree" (Wydick, 1978, p. 738) could be rewritten more concretely to "more attention needs to be paid to the budget".

The prevalence of these linguistic features undermines hinders comprehension of the intended meaning of legal texts and may contribute to inconsistency in the application of the law across cases, which is central to impartiality. When legal language permits multiple interpretations, similar – or even identical – cases are likely to be treated differently. Because impartiality requires that like cases be treated alike (Dworkin 1986, p. 165; Rosthtein and Teorell 2008), this inconsistent treatment represents a breach of impartiality. In other words, impartiality depends not only on excluding factors outside the law stipulations (as emphasized by Rothstein & Teorell 2008), but also on consistently applying the law – something that fundamentally relies on the clarity of legal language.

Figure 1 summarizes our argument: while ambiguity in legal language hinders understanding and expands the potential for divergent interpretations, both interpretation and comprehension influence the consistency with which the law is applied – an essential dimension of impartiality.

Figure 1: Theoretical Argument



Based on the discussion above, we put forward the following hypothesis:

H1: Linguistic ambiguity of legal texts diminishes the consistency with which the law is applied.

4. Data and Method

To test the hypothesis, we conducted an online population-based survey experiment using a two-group, between-subjects design. The survey participants are current and former government officials from around the world. We investigated whether the clarity of legal language influences the consistency with which survey participants interpret and apply the law by randomly assigning them to either a clear or an unclear version of the same legal text on zoning regulations. To this end, the participants were presented with a vignette and asked to assume the role of case officers responsible for issuing building permits on behalf of a local government. They received background information, an excerpt from the law, and a description of a specific case for which they had to determine the appropriate outcome.

We collected responses anonymously from 920 individuals who voluntarily participated in an online survey via a crowdsourcing platform Prolific that outperforms competitors like Amazon Mechanical Turk in terms of the share of participants who meaningfully respond to the researcher's question (Douglas et al. 2023; Peer et al. 2021). Appendix B details platform selection.

Participants were randomly assigned to either a treatment or control group. Both groups received a substantively identical legal text, containing the same conditional elements. The experimental manipulation involved varying the linguistic style: the treatment group received a version written in a traditional legal style featuring linguistic elements known to hinder comprehension and

potentially affect judgment consistency, while the control group received a plain language version aimed at enhancing understanding and minimizing divergent interpretations.

Prolific provides access to a high-quality global participant (Douglas et al., 2023; Peer et al., 2021) and ensures that respondents receive fair monetary compensation. The full questionnaire, implemented in Qualtrics, can be found in Appendix A.

The study was pre-registered on March 28, 2025.[2] Data collection began on April 8, 2025, at 10 AM, and the final response was recorded on April 9th, 2025 at 11 AM. The average time to complete the survey was 5 minutes and 35 seconds. All responses in the dataset are complete and 920 participants were compensated at an hourly rate of £6.55 (i.e., on average £0.5 per response). Appendix B discusses the procedure in detail.

4.1. Population and Sample

Since the study sought to examine decision-making within government organizations, we targeted individuals with prior experience in government or public administration. Any registered Prolific user worldwide with such experience was eligible to participate. Participants were recruited from Prolific's pre-screened pool by selecting the category "Past Experience in Government & Public Administration".

To ensure sufficient power to detect even small effects, the target sample size was based on the assumed small effect size (Cohen's $d \leq 0.2$). Due to uncertainty in the expected effect size – given the lack of clear indications from prior studies – the effect size analysis recommended a minimum of 800 respondents (400 per group). Anticipating some data may not meet quality standards, we set the target sample size at 1,000.

To enhance data quality, the following exclusion criteria were applied. First, incomplete, invalid, or returned responses were excluded. Second, participants were required to complete a reCAPTCHA test, which helps detect and filter out automated responses (i.e., bots) (Kennedy et al. 2020). Responses with a reCAPTCHA score below 0.5 were excluded from the dataset. Third, an attention check was administered to screen out participants who fail to answer the question

[2] The registration DOI at the Open Science Framework is https://doi.org/10.17605/OSF.IO/SQ8XJ.

correctly, as such responses suggest a high likelihood of inattentiveness or random answering. In line with Prolific's policy, these participants were excluded and not compensated (Prolific 2025b).

4.2. Experimental Design

Both groups received the same background scenario: they were asked to assume the role of a case officer working for the planning office in "Town A", responsible for reviewing building permit applications and ensuring regulatory compliance. The scenario focused on the construction sector. Although many respondents may have had little or no prior knowledge in this area, the task required no technical expertise and was devoid of specialized terminology.

They were then presented with an excerpt from a realistic but invented zoning law – *The Zoning Regulation* (Act 2025:1742). The experimental manipulation occurred at this stage: the legal text differed in linguistic complexity between the two groups, while information was identical.

After reading the law, both groups received the same brief case description of a building permit application to construct an attic apartment. Participants were asked to resolve the case by selecting one of four response options, only one of which was correct. This design made it possible to assess whether linguistic complexity in legal texts leads to greater variation in decision-making.

The experiment was based on a real-life building permit application for constructing an attic apartment in central Gothenburg, Sweden. We obtained all relevant documents from the Gothenburg municipality, including national and local laws, the application and the responses of city planning office. The examination of these materials allowed us to gain a detailed understanding of the applicable laws and the building permit procedure, which was then transferred into the experimental design to ensure its validity. Appendix C discusses experimental design in detail.

4.2.1. Legal Text

The two versions of the text convey the same substantive content – requirements for a building permit – but differ in legal linguistics. The legal requirements reflect the provisions from the

Swedish Planning and Building Act (SFS 2010:900) and the relevant Swedish administrative authority's (Boverket) Building Regulations (BBR, 2011, 3:1 & 3:2). We chose to include four requirements that are easy to understand even if participants are unfamiliar with the construction sector. They are a full list of *documents, heritage approval, accessibility provision,* and a final paragraph describing *conditions* for the building permit and when the construction can begin.

The treatment is binary, with participants assigned either to the treatment (1) or the control group (0). The treatment group received a version of the legal text incorporating seven linguistic features – discussed in the theory section – known to influence comprehension and interpretation of legal texts. These include passive voice, negative constructions, nominalizations, syntactic embedding, a Latin expression, and ambiguous formulations and pronoun placement. Care was taken not to overload the text with too many comprehension-inhibiting features to prevent participants from disengaging and responding at random. In contrast, the control group is presented with a clear version of the legal text, which explains the regulatory requirements in language free from these problematic features. Appendix C provides a detailed explanation of the manipulation, including permit requirements and the linguistic elements.

4.2.2. Response Options

Although real-life decisions encompass many possibilities, the number of response options was restricted to keep the experiment manageable while still permitting meaningful measurement of variance. Furthermore, the number of response options needed to strike a balance: the correct choice should not be made too obvious, yet an excessive number of alternatives could introduce undue complexity, given the survey's genre constraints and potential difficulties for participants in interpreting the case. After evaluating several alternatives (see Table 2 in Appendix C for details), we concluded that four response options offered an optimal balance. The response options – "Pending—request revision", "Approve", "Reject", and "Approve with an exemption" – were designed to be equally plausible, and pilot participants (N = 40) selected them at roughly equal rates. This alleviated concerns that selection differences would reflect inherent option biases rather than the experimental manipulation. To further mitigate potential response bias, the order of the response options was randomized.

4.3. Sample Demographics

Socio-demographic and background variables were collected to characterize participants and enable control for individual differences between the groups. Some variables were derived directly from Prolific's prescreening data (e.g., gender, age, first language, and country of origin), while others were obtained through the survey (e.g., legal background, education level, and government employment). These questions are placed at the beginning of the study, prior to the experimental block.

All participants were required to have proficient English language skills, so a dummy variable is created to distinguish between participants for whom English is the first language and those for whom it is not.

Gender is categorized into three groups: "Male," "Female," and "Prefer not to say". Age is a continuous variable that ranges from 18 to 99 (the minimum age of participants is 18 on Prolific), then recoded into a 4-point categorical variable. Education is measured on a 7-point scale and recoded into three categories: "Low", "Intermediate", and "High". Additionally, a variable indicating whether participants have completed any formal legal education is included; it is reported as a dummy, either as "Yes, I have taken at least one law course" or "No, I have never taken a law course". Furthermore, the government position variable is also binary, indicated as either "Currently working in government" or "Previously working in government". Lastly, participants' country of origin is included, and is based on Prolific's data.

Prolific collects approval ratings, reflecting the participants' experiences using Prolific and their approval levels in previous studies. The more studies a participant has been approved for, the higher the score (Prolific, 2025c). We classify this variable as follows: new to moderate experience, ranging from 0 to 100; experienced participants, from 100 to 1000; and highly experienced participants, 1000 and above.

Finaly, the survey features a multiple-choice test to assess reading comprehension and two manipulation checks (see Appendix A and C).

5. Analysis

5.1. Descriptive Statistics

After applying the exclusion criteria, the final dataset includes 885 respondents, with 445 in the treatment group and 440 in the control group. Table 1 reports descriptive statistics. No systematic differences were observed between groups (see Table 2 and 3 in Appendix D), indicating successful randomization.

49% of the treatment group and 59% of the control group selected the correct response option (Pending – request revision). 26% of the treatment group and 24% of the control group chose "Approve the application". 9% of the treatment group and 7% of the control group chose "Reject". 16% of the treatment group and 10% of the control group selected "Approve with an exemption".

Manipulation checks confirmed the treatment effect: 88% in the treatment group and 90% in the control group answered the objective question correctly. On the subjective clarity rating, the treatment group rated the text lower ($M = 6.49$, $SD = 2.37$) than the control group ($M = 7.98$, $SD = 1.69$).

Table 1: Descriptive Statistics

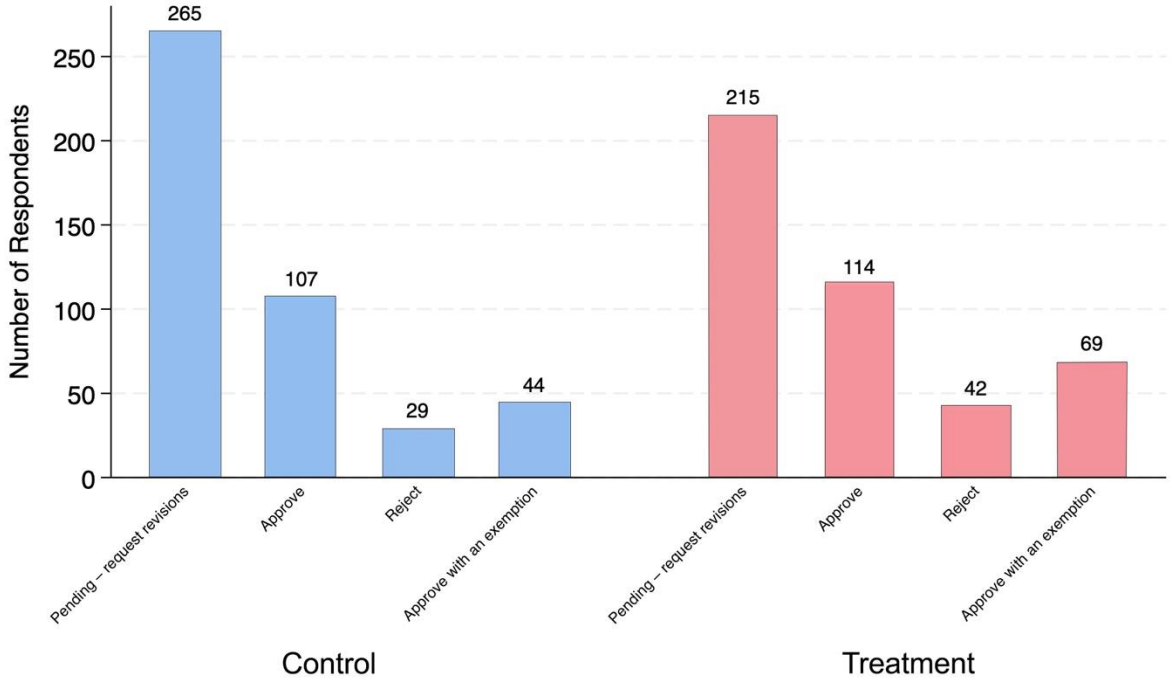
Variable	obs	mean	SD	min	max
Treatment Group					
Decision-Making	440	1.92	1.10	1	4
Gender	440	1.46	0.50	1	3
Age	440	40.56	13.67	19	87
Education	440	2.81	0.39	2	3
Legal Education	440	1.58	0.49	1	2
Government experience	440	1.58	0.49	1	2
English first language	440	0.79	0.41	0	1
Approval rating	440	1.99	0.69	1	3
Reading comprehension	440	0.82	0.39	0	1
Objective manipulation check	440	0.88	0.32	0	1
Subjective manipulation check	440	6.49	2.37	1	10
Control Group					
Decision-Making	445	1.67	0.97	1	4
Gender	445	1.44	0.51	1	3
Age	445	41.52	13.88	18	81
Education	445	2.82	0.39	2	3
Legal Education	445	1.62	0.49	1	2
Government experience	445	1.61	0.49	1	2
English first language	445	0.85	0.35	0	1
Approval rating	445	1.98	0.71	1	3
Reading comprehension	445	0.85	0.36	0	1
Objective manipulation check	445	0.90	0.30	0	1
Subjective manipulation check	445	7.98	1.69	2	10

5.2. Analysis of Variance Across Groups

We examine the effect of law clarity on consistency in decision-making by testing for differences in variance across the groups. More specifically, we evaluate whether participants in the treatment group, presented with the less clear law, demonstrate greater variability in their decision-making compared to the control group, who received the clearer version of the law.

Graph 1 visualizes the distribution of answers between clear (Control) and non-clear (Treatment) language groups. The difference in variance can be discerned visually: the answers are more spread out in the treatment group, suggesting greater inconsistency.

Graph 1: Distribution of Answer Alternatives



To assess whether the distribution of responses systematically differs between groups, we employ a Chi-square test for homogeneity, designed to compare the distribution of categorical variables across two or more groups (Agresti, 2008, pp. 225–239; StataCorp, 2023). Since a Chi-square test for homogeneity computes differences in frequencies between categories, normality in distribution is not a prerequisite, unlike when working with continuous data. The Chi-square test requires each category to hold at least five observations (Agresti, 2008, p. 228), a condition fulfilled by this sample.

Following Agresti (2008, p. 225), the test is computed as:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

O denotes the observed frequencies, and *E* is the expected frequencies calculated under the assumption that there is no difference in the distribution of outcomes across groups. To reject the null hypothesis and to conclude that the distributions differ, the test statistic must be associated with a p-value below the conventional statistical significance threshold of 5% ($p < 0.05$).

Table 2: Chi-squared Test for Homogeneity: Comparison of Decision Consistency Across Groups

Group	*Decision Alternatives*				
	Pending—request revision	**Approve**	**Reject**	**Approve with exemption**	Total
Control Group	265 (54%)	107 (48%)	29 (41%)	44 (39%)	445 (100%)
Treatment Group	215 (46%)	114 (52%)	42 (59%)	69 (61%)	440 (100%)
Total	480 (54%)	221 (25%)	71 (8%)	113 (13%)	885

Chi-Squared	13.31
Degrees of Freedom	3
p-value	0.004

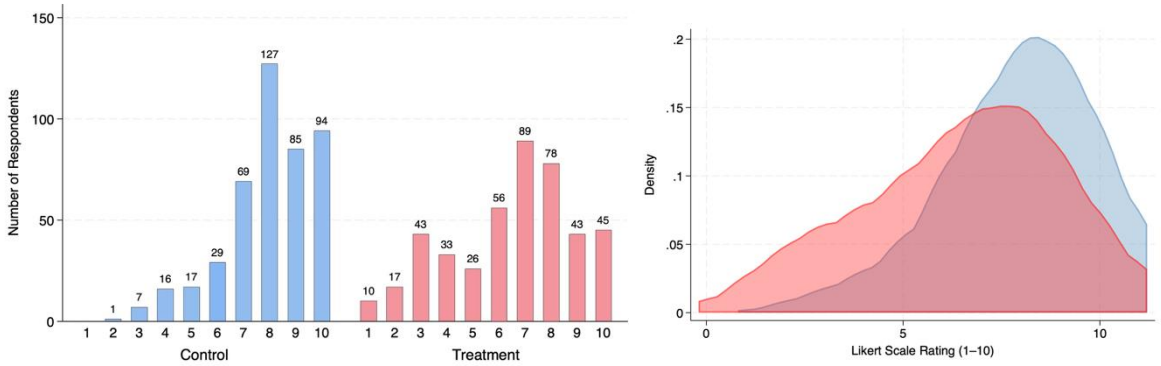
Note. Frequency distributions are presented for each category, and the corresponding percentage within parentheses.

As shown in Table 2, the Chi-squared is statistically significant ($p < 0.004$), indicating a systematic difference in variance across the groups. These results support our hypothesis that reduced linguistic clarity in legal texts is associated with less consistent decision-making.

5.3. Regression Analysis

In addition to analyzing variance in decision-making, we also examine the determinants of participants perceived clarity of the legal texts. The experiment participants were asked to rate the clarity of the legal text on a Likert scale from 0 to 10. Graph 2 depicts the distribution of the responses by group, showing that participants in control group (blue) perceived their legal text as clearer, compared to the treatment group (red), thereby providing addition support to our hypothesis.

Graph 2: Perceived Clarity of Legal Text



We perform an OLS regression (Table 3) testing whether the variable *Treatment*, where unclear law = 1 and clear law = 0, influences the participants perception of clarity, controlling for a set of individual-level characteristics.

Table 3: Result of OLS Regression

	1	**2**	**3**	**4**
Treatment	-1.49***	-1.49***	-1.53***	-1.5***
	(0.14)	(0.14)	(0.14)	(0.14)
Age		-0.07	-0.01	0.07
		(0.08)	(0.09)	(0.09)
Gender		0.18	0.17	0.23
		(0.14)	(0.14)	(0.14)
Education		0.27	0.09	0.09
		(0.18)	(0.18)	(0.18)
Legal education			0.61***	0.49***
			(0.14)	(0.15)
Government position			-0.11	-0.11
			(0.14)	(0.14)
English first language			-0.18	-0.15
			(0.18)	(0.18)
Approval rating				-0.45***
				(0.11)
Reading comprehension check				-0.14
				(0.19)
Objective manipulation check				0.55*
				(0.23)
Constant	7.98***	7.12***	8.81***	8.91***
	(0.01)	(0.62)	(0.69)	(0.70)
Observations	885	885	885	885
Adjusted R^2	0.12	0.12	0.13	0.15

Note. Standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In Model 1, which only includes the binary treatment and the clarity rating, the coefficient for *Treatment* is statistically significant and negatively signed, suggesting that the less clear text affected the perception of clarity. On average, participants in the treatment group rated the clarity of the legal text approximately 1.5 points lower than those in the control group.

This finding is robust to the introduction of basic demographic variables (Model 2), specific skills that may affect the perception of legal clarity (Model 3), as well as the approval rating, the reading comprehension check, and the objective manipulation check (Model 4). Treatment remains statistically significant at the 99% confidence level and negatively signed across all models. Furthermore, the coefficient size remains stable, with membership in the treatment group associated with a 15% decrease in the perceived text clarity.

It is worth noting that legal education is associated with a higher subjective perception of the clarity of legal text. Additionally, a higher approval rating is linked to a reduced clarity rating, while a correctly answered objective manipulation check is associated with a slightly higher clarity rating.

The results of the standard OLS diagnostic tests suggest that the key assumptions (linearity, homoscedasticity, normality of residuals, no multicollinearity, influential observations) of the OLS analysis are met.

The results of the OLS analysis show that participants in the treatment group, who were presented with the less clear legal text, rated its clarity significantly lower than those in the control group. This finding, combined with the finding from the analysis of variance of decision-making outcomes, suggests that less clear legal texts are more difficult to comprehend and also contributes to greater inconsistency in decision-making.

6. Discussion

In this section, we reflect on potential limitations of the study, including concerns related to cognitive load, generalizability, internal validity, measurement reliability, and representativeness of the sample. First, the treatment group may have been exposed to a higher cognitive load than the control group. This could have led to confusion**,** prompting some participants to disengage and select an answer without fully processing the content. The experiment was carefully

designed to mitigate such risks. On one hand, it needed to incorporate appropriate linguistic features to evaluate their impact on decision-making. On the other hand, it required a concise and focused design to maintain participant attention and ensure decisions were based solely on the legal text and considerations of case descriptions. In real decision-making, legal texts are often complex. If the experimental material were overly simplified, it would fail to reflect real challenges. Therefore, a certain degree of complexity was necessary to create a credible scenario and to enable participants to make as informed decisions as possible given the legal language provided.

To further address concerns about cognitive load, we have simplified the study by reducing both the number of questions and response options (Stockemer and Bordeleau 2019, pp. 38–43). Additionally, we have minimized the condition elements to help participants to focus on the core aspects of the decision-making task. These measures aim to ensure that responses are influenced primarily by the linguistic clarity of the legal text, rather than by extraneous complexity. Both versions of the legal texts contain approximately the same number of words; the key difference lies in the clarity of their presentation. While vignette experiments face criticism for lacking realism and for being improbable (Mutz 2011, p. 59), our experiment was based on a real-life case, which mitigates concerns associated with entirely fictional scenarios.

Moreover, experiments containing hypothetical scenarios can suffer from social desirability bias, when respondents base their decisions on what is considered acceptable or wrong according to social norms rather than on their own beliefs (Mutz 2011, p. 61; Stockemer & Bordeleau 2019, p. 41). However, as our questionnaire avoids value-laden or sensitive topics, we do not expect participants to respond in a socially desirable way.

Including attention checks is standard in survey research, and although concerns exist about potential threats to validity, evidence suggests they do not meaningfully affect results (Kung et al. 2018, p. 264). In this study, the attention check was placed before the experimental manipulation as an additional precaution.

Another query concerns how reading comprehension can be assessed through just one multiple-choice question. It can present a challenge, partly because a single question may be insufficient to reflect participants' abilities; however, more extensive approaches are not feasible due to the

constraints of a five-minute survey. This variable is not part of the main analysis, and if you cannot handle this simple task, you may struggle to understand complicated legal language. Therefore, we believe that it gives an appropriate indication.

For the variable approval rating, new members on Prolific consequently have low ratings, as the variable reflects their experience in answering surveys. It provides a measure on experience on Prolific, but it does not indicate how proficient participants are in decision-making scenarios. Government officials may be experienced in interpreting legal texts but less experienced in answering surveys. Consequently, these individuals will naturally have a low approval rating. Participants with lower ratings may take longer to answer surveys than participants with high ratings. However, this is not a major problem since most participants have high approval ratings.

As a final point, we consider the generalizability of the findings. The project's several features make its findings generalizable to broader bureaucratic decision-making. First, the study employs a carefully designed experiment that captures the dimensions it aims to investigate. Second, the experimental materials are based on real laws and incorporate linguistic elements commonly found in formal legislation. Third, the sample consists of current and former government officials, ensuring that the decision-making context closely mirrors real-world bureaucratic processes. Finally, the case scenario itself is inspired by a real case from the urban planning administration in Gothenburg, and the decisions made in the experiment are grounded in the same legal framework used in that setting.

Although the case pertains to zoning regulation, the linguistic features under examination are not unique to this legal domain. These elements are broadly characteristic of legal language across different sectors. Therefore, the findings are not confined to urban planning but may plausibly extend to other areas of law and administration. Moreover, the participants, all of whom are or have been public officials, were not limited to roles within a specific sector, further reinforcing the broader applicability of the results.

7. Conclusion

This project began with the observation that certain linguistic features used in legal texts can lead to misunderstandings and leave greater room for subjective interpretation, making it more

difficult for implementers to apply the law consistently. Such inconsistency in decision-making poses a threat to the principle of impartiality – one of the foundational pillars of high-quality government (QoG). A review of the existing QoG literature revealed a striking absence of research examining the relationship between linguistic clarity and impartiality. This paper set out to address that gap.

Drawing on the legal linguistics research, we developed a theoretical argument linking language clarity in legal texts with impartiality in bureaucratic decision-making through consistent application of the law. We argued that bad wording of laws can reduce understanding, allowing for divergent interpretations that lead to greater inconsistency in the application of the law, ultimately threatening impartiality.

We then empirically tested this argument through an online survey experiment involving nearly 900 former and current government officials from 33 countries. The participants were divided into two groups, each of which was given a scenario inspired by a real building permit application based on Swedish zoning laws. The experimental treatment incorporated seven linguistic elements to produce a less clear variant of the law. Our results showed that the treatment group displayed greater variability in their decisions, compared to control group, which received a clearer text, thereby supporting our theoretical argument.

By demonstrating that unclear language in legal texts can cause unintended variations in decision-making – and even results in incorrect application of the law, risking unequal treatment of similar cases – this paper highlights the need to broaden the definition of Quality of Government. While Rothstein and Teorell's (2008) definition emphasizes the threat to impartiality from factors not "stipulated in the policy or the law" (p. 170), this paper argues that the clarity with which laws *are* formulated is also important for ensuring impartial implementation.

Integrating legal clarity into the Quality of Government (QoG) framework requires rethinking both the concept's scope and its operationalization. Impartiality has often been operationalized by the absence of discretionary bias or favoritism. Adding language clarity adds a further dimension to impartiality – minimal interpretive divergence across implementers.

Measuring interpretive variance (e.g., through vignettes or linguistic analysis) becomes a legitimate indicator of QoG. Readability scores (e.g., Flesch-Kincaid), counts of complex constructions (nominalizations, embedded clauses), or expert surveys on textual clarity can supplement existing QoG indicators. This enriches cross-national comparisons by capturing variation in statutory legibility, including comparing across legal families. Empirical QoG assessments can incorporate experimental or survey-based modules (as in this paper) to test whether language clarity in legal texts predicts outcomes such as case-processing times, appeal rates, or administrative complaints.

While unclear laws undeniably affect QoG, in countries where public administration is grappling with high levels of corruption, politicization, or nepotism, the factors not "stipulated in the policy or the law" (Rothstein & Teorell 2008, p. 170) may carry greater weight than the wording of legal provisions. In other words, the effect of legal clarity may be more pronounced in countries where the quality of government is already relatively high. This conjecture warrants empirical investigation in future research.

To support impartial implementation, policymakers should prioritize the simplification of legal language. Unclear laws increase the need for legal support and complicate the consistent application of rules. As Fernández-i-Marín et al. (2024b) stress, it is essential to align the regulatory burdens placed on agencies with the resources available to them. Simplifying legal language can reduce this burden and promote more consistent and impartial decision-making. Policymakers could adopt plain-language drafting manuals, statutory templates, and regular legislative "health checks" as part of a broader toolkit to enhance QoG. Recognizing that clarity and comprehension are closely linked, governments might also invest in ongoing training for drafters, judges, and administrators on linguistic best practices, complementing textual reform with capacity building.

As the paper demonstrates, individuals with formal legal education are better equipped to deal with unclear legal texts than those without such training. This finding suggests that legal literacy could be an important criterion for recruitment into the public bureaucracy. However, before advancing this as a policy recommendation, further research is needed to evaluate the potential benefits of integrating legal literacy into standard merit-based recruitment practices.

This paper introduces a novel lens on Quality of Government by unpacking how clarity of legal language shapes the impartiality of its implementation, thereby deepening our understanding of the institutional foundations of good governance.

References

Agnafors, M. (2013). Quality of Government: Toward a More Complex Definition. *American Political Science Review, 107*(3), 433–445.

Agresti, A. (2008). *Statistical Methods for the Social Sciences* (4th ed.). Pearson.

Ahlerup, P., Baskaran, T., & Bigsten, A. (2016). Government Impartiality and Sustained Growth in Sub-Saharan Africa. *World Development, 83*, 54–69.

Blasie, M. (2022). The Rise of Plain Language Laws, University of Michigan Law Review 76(2), 447-524.

Boverket. (2011). Boverket's Building Regulations (BBR), 3:1 & 3:2. Retrieved 2025-04-07 from https://www.boverket.se/en/start/

Charron, N., Dahlström, C., & Lapuente, V. (2012). No law without a state. *Journal of Comparative Economics, 40*(2), 176–193.

Charrow, R. P., & Charrow, V. R. (1979). Making Legal Language Understandable: A Psycholinguistic Study of Jury Instructions. *Columbia Law Review, 79*(7), 1306-1374.

Dahlberg, S., & Holmberg, S. (2014). Democracy and Bureaucracy: How Their Quality Matters for Popular Satisfaction. *West European Politics, 37*(3), 515–537.

Dahlström, C., & Lapuente, V. (2017). *Organizing Leviathan: Politicians, Bureaucrats and the Making of Good Government*. Cambridge: Cambridge University Press.

Dale, W. (1977). *Legislative Drafting: A New Approach; A Comparative Study of Methods in France, Germany, Sweden and the United Kingdom*. London: Butterworth.

Diamond, S. S., & Levi, J. N. (1996). *Improving decisions on death by revising and testing jury instructions. Judicature 79*(5), 224-232.

Douglas, B. D., Ewell, P. J., & Brauer, M. (2023). Data quality in online human-subjects research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. *PLOS ONE, 18*(3), e0279720.

Duvanova, D. (2017). Delegation of Political Authority to Bureaucratic Agencies in an Authoritarian Regime: An Analysis of Regional Economic Regulation in Kazakhstan. *Studies in Comparative International Development, 52*(1), 87–114.

Dworkin, R. (1986). *Law's Empire*. The Belknap Press.

Esaiasson, P., Gilljam, M., Oscarsson, H., Sundell, A., Towns, A., & Wängnerud, L. (2024). *Metodpraktikan*. Gothenburg: Gothenburg University Press.

Etikprövningsmyndigheten. (2023). *Guide to the Ethical Review of Research on Humans*. https://etikprovningsmyndigheten.se/en/

Fernández-i-Marín, X., Hinterleitner, M., Knill, C., & Steinebach, Y. (2024a). Bureaucratic overburdening in advanced democracies. *Public Administration Review, 84*(4), 696–709.

Fernández-i-Marín, X., Hinterleitner, M., Knill, C., & Steinebach, Y. (2024b). Policy growth, implementation capacities, and the effect on policy performance. *Governance, 37*(3), 927–945.

Fernández-i-Marín, X., Steinbacher, C., Steinebach, Y., & Knill, C. (2024c). Bureaucratic Quality and the Gap between Implementation Burden and Administrative Capacities. *American Political Science Review, 118*(3), 1240–1260.

Gerber, A. S., Huber, G. A., Doherty, D., & Dowling, C. M. (2016). Why People Vote: Estimating the Social Returns to Voting. *British Journal of Political Science, 46*(2), 241–264.

Gibbons, J. (1999). Language and the Law. *Annual Review of Applied Linguistics, 19*, 156–173.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1-76.

Glaeser, E. L., & Shleifer, A. (2002). Legal origins. *The Quarterly Journal of Economics*, *117*(4), 1193-1229.

Gothenburg City Planning Office. (2023a, April 27). *Comments on your application: Case number SBN 2023-002449* (in Swedish). Gothenburg: Gothenburg City Planning Office.

Gothenburg City Planning Office. (2023b, June 20). *Comments on your application: Case number SBN 2023-002449* (in Swedish). Gothenburg: Gothenburg City Planning Office.

Gothenburg City Planning Office. (2023c, November 8). *Decision on start notice: Case number SBN 2023-002449* (in Swedish). Gothenburg: Gothenburg City Planning Office.

Hauser, D. J., Ellsworth, P. C., & Gonzalez, R. (2018). Are Manipulation Checks Necessary? *Frontiers in Psychology*, *9*, 1-10.

Holmberg, S., Rothstein, B., & Nasiritousi, N. (2009). Quality of Government: What You Get. *Annual Review of Political Science*, *12*(1), 135–161.

Huber, G. A., Hill, S. J., & Lenz, G. S. (2012). Sources of bias in retrospective decision making: Experimental evidence on voters' limitations in controlling incumbents. American Political Science Review, 106(4), 720-741.

Kennedy, R., Clifford, S., Burleigh, T., Waggoner, P. D., Jewell, R., & Winter, N. J. G. (2020). The shape of and solutions to the MTurk quality crisis. *Political Science Research and Methods*, *8*(4), 614–629.

Kimble, J. (1998). The Great Myth that Plain Language is not Precise, *Scribes Journal of Legal Writing* 7, 109-118.

Kimble, J. (1994). Answering the Critics of Plain Language, *Scribes Journal of Legal Writing* 5, 51-86.

Kimble, J. (1992). Plain English: A Charter for Clear Writing, *Thomas M. Cooley Law Review 9*(1), 1-58.

Kolvani, P., & Nistotskaya, M. (2025). On mechanisms of meritocratic recruitment: Competence and impartiality. *Journal of Institutional Economics*, *21*, e12.

Kung, F. Y. H., Kwok, N., & Brown, D. J. (2018). Are Attention Check Questions a Threat to Scale Validity? *Applied Psychology*, *67*(2), 264–283.

La Porta, R. L., Lopez-de-Silanes, F., & Shleifer, A. (2008). The Economic Consequences of Legal Origins. *Journal of Economic Literature*, *46*(2), 285–332.

La Porta, R., Vishny, R., Florencio, L.-S., & Shleifer, A. (1999). The Quality of Government. *Journal of Law, Economics, and Organization*, *15.1*, 222–279.

Majambere, E. (2011). Clarity, precision and unambiguity: Aspects for effective legislative drafting. *Commonwealth Law Bulletin*, *37*(3), 417–426.

Milosav, Đ., & Nistotskaya, M. (2024). Unpacking the relevance of interpersonal trust in the blockchain era: Theory and experimental evidence. *Journal of Behavioral and Experimental Finance*, *42*, 100901.

Mousmouti, M. (2012). Operationalising Quality of Legislation through the Effectiveness Test. *Legisprudence*, *6*(2), 191–205.

Mutz, D. C. (2011). *Population-based survey experiments*. Princeton University Press.

Nistotskaya, M. (2020). Quality of Government (QoG) as Impartiality: Review of the literature on the causes and consequences of QoG. *KIPA Public Policy Review*, *1*(1), 25–49.

Nistotskaya, M., Charron, N., & Lapuente, V. (2015). The wealth of regions: Quality of government and SMEs in 172 European regions. *Environment and Planning C: Government and Policy*, *33*(3), 1125–1155.

Nistotskaya, M., & Cingolani, L. (2016). Bureaucratic Structure, Regulatory Quality, and Entrepreneurship in a Comparative Perspective: Cross-Sectional and Panel Data Evidence. *Journal of Public Administration Research and Theory*, *26*(3), 519–534.

Olander, P. (2021). Economic diversification, homogeneity of interests, and the impartiality of government. In A. Bågenholm, M. Bauhr, M. Grimes, & B. Rothstein (Eds.), *The Oxford handbook of the quality of government* (pp. 378–398). Oxford University Press.

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27.

Pearson, P. D., & Hamm, D. N. (2005). *The Assessment of Reading Comprehension: A Review of Practices—Past, Present, and Future*.

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2021). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, *54*(4), 1643–1662.

Plan- och bygglagen (SFS 2010:900). *Plan- och bygglag* (PBL). Stockholm: Regeringskansliet.

Posner, R. A. [1973] 2007. *The Economic Analysis of Law.* 7th ed. New York: Aspen Publishers.

Prolific. (2025a, April 24). *Who are the participants on Prolific? Prolific*.

Prolific. (2025b, April 24). *Prolific's payment principles. Prolific*.

Prolific. (2025c, April 24). *How do I use Prolific's demographic pre-screening? Prolific*.

Randall, J. (2014). Tackling "Legalese": How Linguistics Can Simplify Legal Language and Increase Access to Justice. In J. Emonds & M. Janebová, *Language Use and Linguistic Structure* (1st ed., pp. 239–254). Palacký University Press.

Rothstein, B., & Teorell, J. (2008). What Is Quality of Government? A Theory of Impartial Government Institutions. *Governance*, *21*(2), 165–190.

Schriver, K.A. (2017). Plain Language in the US Gains Momentum: 1940–2015, *IEEE Transactions on Professional Communication*, *60*(4), 343-383.

Shohamy, E. (1984). Does the testing method make a difference? The case of reading comprehension. *Language Testing*, *1*(2), 147–170.

Solan, L. M. (1995). Judicial Decisions and Linguistics Analysis: Is there a linguist in the court? *Washington University Law Quarterly 73*(3), 1069-1080.

Sparling, R. A. (2018). Impartiality and the Definition of Corruption. *Political Studies*, *66*(2), 376–391.

StataCorp. (2023). Stata Base Reference Manual Release 19. *College Station: Stata Press.*

Stockemer, D., & Bordeleau, J.-N. (2019). *Quantitative Methods for the Social Sciences: A Practical Introduction with Examples in SPSS and Stata*. Springer International Publishing.

Tiersma, P. (1999). *Legal Language*. Chicago: University of Chicago Press.

Thomas, R. (1985). Plain English and the Law, *Statute Law Review*, 6, 139-151.

U.S Department of Labor. (n.d.). *Minimum Wage*. DOL. Retrieved 17 April 2025, from https://www.dol.gov/general/topic/wages/minimumwage

Vanterpool, V. (2007). A critical look at achieving quality in legislation. *European Journal of Law Reform*, 9(2), 167-204.

Voermans, W. (2011). Styles of Legislation and Their Effects. *Statute Law Review, 32*(1), 38–53.

Williams, C. (2004) *Tradition and Change in Legal English: Verbal Constrictions in Prescriptive Texts*. Oxford: Peter Lang.

Wydick, R. C. (1978). Plain English for Lawyers. *California Law Review, 66*(4), 727–766.

Özdemir, E.Ç. & Akyol, H. (2019). The Development of a Reading Comprehension Test. *Universal Journal of Educational Research, 7*(2), 563–570.

Appendices

Appendix A. The Survey

This appendix reports the full survey questionnaire in the format and order (except for randomized response options) in which it was presented to participants.

Introduction

Survey information

Welcome to the survey!

The Quality of Government Institute at the University of Gothenburg in Sweden is conducting a study on how people experience legal texts. In this survey, you will be randomly assigned to either a "easy" or a "difficult" text and then asked a few questions about your perception of it. The University of Gothenburg complies with the EU General Data Protection Regulation (GDPR). This means that your personal information will be used solely for social science research, and your data and responses will not be shared with third parties. If you have any questions about the study or the survey, please feel free to contact the project's PI Joakim Nilsson at: gusniljofn@student.gu.se

---

The survey will take between 4-6 minutes to complete. Your responses will remain completely anonymous. We appreciate your time and participation! (consent)

Do you agree to participate in this survey?

- ○ Yes
- ○ No (This ends the survey)

End of Block

Start of Block

Please verify (ReCAPTCHA, bots capture)

End of Block

Start of Block

What is the highest level of education you have completed? (Education)

- ○ No formal Education
- ○ Elementary school or equivalent
- ○ High school or equivalent
- ○ College / vocational training
- ○ Bachelor or equivalent
- ○ Master or equivalent
- ○ Doctoral or equivalent

Have you received any formal education in law? (legal education)

- ○ Yes, I have taken at least one course
- ○ No, I have never received formal education in law

Which of the following best describes you? (government position)

- ○ Currently working in Government & Public Administration
- ○ Previously worked in Government & Public Administration
- ○ Have never worked in Government & Public Administration

End of Block

Start of Block

Please select 'Strongly agree' to show you are paying attention to this question (attention check)

- ○ Strongly agree
- ○ Agree
- ○ Disagree
- ○ Strongly disagree

End of Block

Start of Block

Choose the right term for this scenario **(**reading comprehension)

Most elephants are grey, but occasionally, a white elephant is born. In the past, these rare animals were of high spiritual value in India. When one was discovered, it was given to the ruler, who had to care for the elephant despite its limited practical use. Today, when someone gives a gift that is hard to take care of and difficult to give away, we speak of it as

- ○ A grey elephant
- ○ A valuable gift
- ○ An Indian ruler
- ○ A white elephant

End of Block

Start of Block

The construction of new buildings and modification to existing ones are governed by zoning laws, typically established by local governments, which regulate land use and specify the types of structures allowed in different areas. In Town A, companies seeking to construct a new building must apply for a building permit. Their application must demonstrate that the proposed structure complies with zoning regulations. Imagine you are a case officer in Town A's Planning Office, tasked with reviewing and deciding on building permit applications. Please read an excerpt from Town A’s zoning law and the case description, then select the most appropriate decision.

Treatment group

Act (2025:1742) – The Zoning Regulation

§1 *A building permit shall not be issued except upon the submission of the complete requisite architectural renderings, including but not limited to a floor plan, section drawing, and elevation drawing. Any application deficient of the aforementioned or nonconforming to prescribed standards shall be deemed ipso facto incomplete and subject to revision.*

Modification of zoning-regulated structures in densely built areas is not admissible absent final determination by the duly authorized preservation authorities.

Authorization for residential units exceeding 35 square meters in spatial extent shall be precluded absent a duly substantiated demonstration of their accessibility for persons with disabilities, as deemed satisfactory by the competent authorities.

Construction shall not commence absent the prior issuance of the requisite building permit.

Case Description

A company has applied for a building permit to convert an unused attic into an apartment in a building, located in a densely built area covered by a zoning plan. The submitted documents include a floor plan and a section drawing, showing that the attic is 40 square meters and is accessible for people with disabilities.

Which of the following best describes your decision on the case? Select one: (DV)

- ○ Pending – request revisions
- ○ Approve
- ○ Reject
- ○ Approve with an exemption

Control group

Control Act (2025:1742) – The Zoning Regulation

§1 *To obtain a building permit, the application must meet the following requirements: The application must include a floor plan, section drawing, and elevation drawing. If any required documents are missing, the application shall be revised and resubmitted.*

If the building is in a densely built area covered by a zoning plan, heritage conservation approval is required.

Apartments larger than 35 square meters must be accessible for people with disabilities.

Construction cannot start until a construction permit has been issued.

Case Description

A company has applied for a building permit to convert an unused attic into an apartment in a building, located in a densely built area covered by a zoning plan. The submitted documents include a floor plan and a section drawing, showing that the attic is 40 square meters and is accessible for people with disabilities.

Which of the following best describes your decision on the case? Select one: (DV)

- ○ Pending – request revisions
- ○ Approve
- ○ Reject
- ○ Approve with an exemption

End of Block

Start of Block

Which of the following must be included in the application? (objective manipulation check)

- ○ The application must include a floor plan, section drawing, and elevation drawing
- ○ The application must include a personal statement from the applicant
- ○ The applicant must submit a financial plan for the construction project
- ○ The application is automatically accepted as long as a zoning plan exists

End of Block

Start of Block

How would you rate the clarity of the legal text on which you based your decision? (subjective manipulation check)

	Not at all clear 1	2	3	4	5	6	7	8	9	Extremely clear 10
	○	○	○	○	○	○	○	○	○	○

End of Block

Start of Block

What is your Prolific ID? Please note that this response should auto-fill with the correct ID

_____

End of Block

Appendix B. Online Survey

Selecting Online Platform

Conducting survey-experiments and recruiting participants through crowdsourcing platforms is a relatively new phenomenon that, in recent years, has gained increased popularity among researchers (Milosav & Nistotskaya 2024; Kennedy et al. 2020; Palan & Schitter 2018). Several prominent social science studies have conducted survey-experiments using such platforms (e.g., Gerber et al. 2016; Huber et al. 2012). Using crowdsourcing platforms to conduct survey-experiments holds many benefits, not least in terms of convenience and cost-effectiveness (Peer et al. 2021, p. 1643). The use of crowdsourcing platforms has proven to be a reliable method, as demonstrated by numerous successful replications of prior experimental studies (Palan & Schitter 2018, p. 22).

Among crowdsourcing platforms, Amazon Mechanical Turk (MTurk) is the most well-known (Kennedy et al. 2020; Palan & Schitter 2018). However, as indicated by recent reports, MTurk has faced increasing criticism for poor data quality, including frequent use of bots to answer surveys (Kennedy et al. 2020); inattentive responding by participants; and issues related to underpayment, raising both ethical concerns and problems with participants' motivation (Palan & Schitter 2018, p. 23). I chose Prolific as the study's platform, as it emerges a viable alternative; it is a bit more expensive than MTurk, but in contrast, studies show that Prolific offers high-quality data compared to competitors (Douglas et al. 2023; Peer et al. 2021).

Procedure

The study was conducted as follows: Prolific's website presents a list of potential studies to participants who meet the pre-specified eligibility criteria. A title and a brief description are displayed for each study, along with the estimated completion time and the remuneration amount. Participants who choose to take part in the study are automatically directed to Qualtrics, where the survey is administered. Participants can only complete the survey once, and upon completion, they are automatically returned to Prolific, where they receive the payment. Through Qualtrics, the questionnaire's layout, style, and format are adapted for mobile devices and computers, making the survey accessible to a wider audience. After reading the law, both groups

receive the same brief case summary concerning a building permit application to construct an attic apartment. Participants are then asked to resolve the case by selecting one of four response options, only one of which was correct (Appendix C discusses experimental design in details). This design allows us to examine whether exposure to ambiguous language increases variability in decision-making.

The questionnaire is organized as follows. The introductory page provides a brief overview of the survey and its subject matter, and participants must actively consent to participate. Following this, a demographics section and covariates, such as age, education level, legal background, and a comprehension check, are included. The third block contains an attention check, while the fourth consists of a reading comprehension test. After this, participants are randomly assigned to either the treatment or the control group. The treatment group receives a version containing seven linguistic elements discussed in the theory section (see Table 1 in Appendix C for details). The control group receives a plain-language version, which serves as a basis for comparison. The labels ‘treatment’ and ‘control’ are used for clarity, even though both groups are exposed to a legal text.

To prevent skewed group distributions, participants were randomly assigned to the control or treatment group after completing the initial section and immediately before beginning the experimental block. This approach reduces the likelihood that early dropouts get assigned to a group, thereby mitigating attrition-related imbalances.

After reading either the ambiguous or the clearer version of the same law, both groups were presented with an identical case description concerning a building permit application for an attic apartment. Participants were then asked to resolve the case by selecting one of four response options. Appendix C discusses research design in details.

Compensation for Participating

Providing fair compensation is essential not only for meeting ethical standards but also for motivating participants and enhancing data quality (Prolific 2025b). Payment size can influence if participants choose to opt into the study, previous studies imply that data quality is generally

not affected by the level of compensation (Buhrmester et al. 2011, cited in Palan & Schitter, 2018, p. 23). Nonetheless, appropriate compensation remains important, given that the survey arguably entails a higher cognitive load than typical questionnaires. Participants received an average reward of £6.55 per hour, which exceeds Prolific's minimum wage policy of £6 per hour, mirroring the US minimum wage of $7.25 (i.e., about £5.4) (Prolific 2025b; U.S Department of labor n.d.). The actual pay is proportional to the time spent on answering the questions and should not be excessively high to attract participants who engage in the study solely for monetary reasons (Etikprövningsmyndigheten 2023, p. 33). This study was funded (100%) by the Quality of Government Institute at the University of Gothenburg, and the approximate total cost of carrying out the experiment was £700.

Ethical Consideration

Participants' responses were collected anonymously, and personal data cannot be linked to any individual. The participants were fairly compensated for taking the survey. They were introduced to the topic before they chose to move on from Prolific; they were briefed on the study's purpose, estimated completion time, compensation level, and hourly wage, and were also encouraged to reach out with any queries. The introduction states that the survey was conducted by the Quality of Government Institute at the University of Gothenburg, and participants were informed that the project follows the EU General Data Protection Regulation (GDPR) (Etikprövningsmyndigheten 2023, pp. 86–89), all collected data will only be used for social science research and Individual responses will not be shared with a third party. Participants must actively consent to proceed; if they do not, the survey closes. Throughout the survey, neutral language is employed. We do not consider the subject to be sensitive: neither political affiliation, corruption, political behavior, nor any other related subject was investigated.

Appendix C. Experimental Design

Treatment: Ambiguous and Clearer Legal Texts

Participants were exposed to two different linguistic versions of the substantively identical law – SFS 2010: 900. The first paragraph of the law sets out the *documentation requirements*, including a floor plan, section drawing, and an elevation drawing. The treatment includes the use of passive voice and negative constructions (e.g., "shall not be issued"), nominalization (e.g., "nonconforming"), syntactic embedding ("including but not limited to a floorplan"), and a Latin expression ("ipso facto"). The second paragraph states that if the building is situated in a densely constructed area covered by a zoning plan, a *heritage conservation approval* is required. It features passive voice, negative constructions (e.g., "is not admissible"), and nominalization (e.g., "determination"). Third, the *accessibility provision* requirement states that apartments larger than 35 square meters must be accessible for people with disabilities. Here, the treatment text includes passive voice (e.g. "shall be precluded"), nominalization (e.g. "demonstration"), ambiguous pronoun placement (e.g., "their", which could refer either to 'residential units' or 'persons with disabilities'), and ambiguous phrasing (e.g., "as deemed satisfactory by the authorities"). The fourth and final paragraph stipulates that *construction may not begin until the planning office has issued a permit*. It features passive voice and a negative formulation ("shall not commence") and a nominalization ("issuance"). These legal provisions establish a foundation for decision-making regarding the case presented. However, the linguistic interventions make it more difficult to discern the necessary requirements in the text presented to the treatment group. Table 1 below summarizes all the linguistic elements and their placement in the legal text.

The Case

The law text was followed by a case description. Positioning the case description on the same page immediately after the legal text ensured that participants had all necessary information at hand, without needing to revisit earlier sections. Participants were then asked to resolve the case by selecting one of four response options, only one of which was correct.

The case is based on a real-life building permit application for constructing an attic apartment in central Gothenburg, Sweden. By requesting all documents from the Gothenburg municipality, we examined the application and the city planning office's responses in detail until the permit was issued. This allowed us to gain a thorough understanding of the local legislation and to map the progression of a case from application to a final approval. The case provided the knowledge and insights necessary for the experimental design.

Box 1. Real-life Building Permit Application Case

The case involved a tenants' association seeking permission to build an apartment in an unused attic space within a densely populated area governed by a zoning plan of the Gothenburg municipality. The case officer's initial response stated that the construction would "cause a distortion" in the area under the zoning law (with reference to SFS 2010:900, 4 kap, 8 §) and required the applicant to submit a heritage approval along with additional specified documents (Gothenburg City Planning Office 2023a, p. 1). After the applicant provided the requested documents, including the heritage approval, the Planning Office responded: "We cannot grant your application, and therefore suggest that you change your application" (Gothenburg City Planning Office 2023b, p. 1). Although the heritage assessment was positive, the city further suggested that the applicant "should take a holistic approach to the roof landscape" (Gothenburg City Planning Office 2023b, p. 1), necessitating another revision. In the third revision, the case officer determined that the proposed plan did not meet the accessibility requirements specified by national law (BBR 2011, 3:223), prompting yet another revision. In November 2023, the applicant received permission to begin preparatory constriction work, with any further work dependent on meeting the remaining requirements (Gothenburg City Planning Office 2023c).

Answer Options

After reading the law and the case, participants were asked to resolve the case. Positioning the case description immediately after the legal text ensured that participants had all necessary information at hand, without needing to revisit earlier sections.

There are four different decisions to choose from, each of which has a relatively equal chance of being selected, but only one representing the correct decision. Table 2 discusses the underlying reasoning for each of the alternative.

During the design phase, we also considered a fifth option: adding a clause to the legal text that would introduce a time period within which the applicant could submit a heritage approval. This option was excluded, as it risked introducing excessively complexity into the decision-making task, likely prompting participants to respond randomly. We also considered including an alternative in which approval could be granted in exchange for a payment to the case officer. This option was discarded both for ethical reasons – it represents a socially undesirable choice – and because it constitutes an act of corruption not grounded in law, given that no allowance for extra fees exists in the legal text.

Reading Comprehension and Manipulation Checks

As reading comprehension influences individuals' ability to understand complex legal texts (Charrow & Charrow 1979; Diamond & Levi 1996; Randall 2014), a a multiple-choice reading comprehension test – adapted from Shohamy (1984, p. 169) – was conducted.[3] While the test assignment is short and lacks nuances, it serves as an approximate measure of comprehension. We also treat it as an indirect validity check as the failure to answer the reading comprehension question correctly may indicate difficulty in understanding the legal text also. The variable is coded as a binary (0 = fail, 1 = pass), and we use it in the OLS analysis of the assessment of the clarity of the legal text by participants.

The following block includes two manipulation checks. The first – objective manipulation check – asks participants to recall essential information from the legal text. This question assesses their understanding of key elements and probes whether participants in the ambiguous condition are less likely to extract the same basic information. Four answer options are provided, only one of which is correct. A subjective manipulation check follows to compliment the objective measure:

[3] Research in educational psychology supports the reliability of multiple-choice formats for measuring reading comprehension (Özdemir & Akyol 2019, p. 567; Pearson & Hamm 2005).

participants rate the clarity of the legal text on a Likert scale from 0 (Not at all clear) to 10 (Extremely clear). Consistent with the literature (Mutz 2011, p. 65), we place the manipulation checks after the treatment to minimize any influence on participants' decision-making.

Table 1. Treatment Explanation

Conditional Element	Condition Inspired from Law	Clear Legal Text (Control Group)	Unclear Legal Text (Treatment Group)	Linguistic Issue
1. Documentation	*SFS 2010:900, 9 kap, 21 §*	To obtain building permit the application must meet the following requirements: The application must include a floor plan, section drawing, and elevation drawing. If any required documents are missing, the application shall be revised and resubmitted.	A building permit shall not be issued except upon the submission of the complete requisite architectural renderings, including but not limited to a floor plan, section drawing, and elevation drawing **(1, 2, 4)**. Any application deficient of the aforementioned or nonconforming to prescribed standards shall be deemed ipso facto incomplete and subject to revision **(1, 3, 5).**	**1.** Passive voice **2.** Negative sentence **3.** Nominalization **4.** Syntactic embedding **5.** Latin Phrase
2. Heritage approval	*SFS 2010:900, 4 kap, 8 §,* *SFS 2010:900, 9 kap, 13 §*	If the building is in a densely built area covered by a zoning plan, heritage conservation approval is required	Modification of zoning-regulated structures in densely built areas is not admissible absent final determination by the duly authorized preservation authorities **(1, 2, 3)**	**1.** Passive voice **2.** Negative sentence **3.** Nominalization
3. Accessibility provision	*SFS 2010:900, 9 kap, 1 §,* *BBR, 3:223*	Apartments larger than 35 square meters must be accessible for people with disabilities.	Authorization for residential units exceeding 35 square meters in spatial extent shall be precluded absent a duly substantiated demonstration of their accessibility for persons with disabilities, as deemed satisfactory by the authorities **(1, 3, 6, 7)**.	**1.** Passive voice **3.** Nominalization **6.** Ambiguous pronoun placement **7.** Ambiguous formulation
4. Conditions	*SFS 2010:900,10 kap, 3 §*	Construction cannot start until a construction permit has been issued.	Construction shall not commence absent the prior issuance **(1, 2, 3)** of the requisite building permit.	**1.** Passive voice **2.** Negative sentence **3.** Nominalization

Table 2. The Rationale for Decision Alternatives

Answer Option	**Interpretation**	**Reasoning**
1. Pending—request revision	*Correct*	The application includes a floor plan and section drawing, but not an elevation drawing, required by the law. Therefore, the case officer should ask the applicant for a revision. The law states it is *"Subject to revision"* (T) and *"Shall be revised"* (C). The linguistic complexity in "Subject to revision" (T) might cause participants to overlook the revision requirement. Additionally, the building is in a *"Densely built area covered by a zoning plan"* (Case), requiring heritage approval, which was not submitted, which requires a revision.
2. Approve	*Incorrect*	The application lacks required documentation, namely an elevation drawing. The treatment version reads: *"Including but not limited to a floor plan, section drawing, and elevation drawing"* (T), embedded in complex language. Passive voice and nominalization like *"Subject to revision"* (T) might cause confusion. Some participants may incorrectly grant approval because the application includes section drawings and notes that the building is "accessible for people with disabilities", without noticing that the required elevation drawing is missing.
3. Reject	*Plausible, but incorrect*	The proposal meets the accessibility requirements, as stated in the case description. However, the treatment version introduces ambiguity through the clause *"as deemed satisfactory by the competent authorities"* (T), which may lead participants in this group to view the application as acceptable. In contrast, the control version states the requirements more clearly. Participants may also misinterpret the missing elevation drawing as grounds for rejection or dismiss the application due to the absence of a heritage approval, even though this document can be submitted later.
4. Approve with an exemption	*Incorrect*	The elevation drawing is missing and cannot be exempted. The law requires all three documents – floor plan, section drawing, and elevation drawing – and the officer must request a revision if any are missing. This alternative may lead participants to overlook or misunderstand the pathway to revision, creating the false impression that an exemption is possible.

Note. (C) indicate the control group, and (T) indicate the treatment group, and (Case) denote the case description text.

Appendix D. Descriptive Statistics

Table 1. Variable Coding and Frequency Distribution of Data by Group (Counts)

Variable and Categories	**Control**	**Treatment**	**Total**
Total	445	440	885
Dependent variable			
Decision-making (DV)			
Pending—request revision (1)	265	215	480
Approve (2)	107	114	221
Reject (3)	29	42	71
Approve with an exemption (4)	44	69	113
Demographical variables			
Gender			
Female (1)	249	239	488
Male (2)	194	200	394
Prefer not to say (3)	2	1	3
Age			
16–29 years	103	105	208
30–49 years	213	215	428
50–64 years	97	95	192
65–99 years	32	25	57
Education			
Low (1)	-	-	-
Intermediate (2)	82	83	165
High (3)	363	357	720
Legal education			
Yes, I have taken at least one law course (1)	171	183	354
No, I have never taken a law course (0)	274	257	531
Government position			
Currently working in government (1)	174	184	358
Previously worked in government (2)	271	256	527
English as first language			
Yes (1)	380	346	726
No (0)	65	94	159
Other Measures			
Approval rating			
New/moderate (0 and 100 approvals)	117	107	224
Experienced (100-1000 approvals)	219	232	451
Highly Experienced (>1000 approvals)	109	101	210
Reading comprehension			
Fail (0)	66	81	147
Pass (1)	379	359	738
Manipulation check (objective)			
Incorrect answer (0)	44	51	95
Correct answer (1)	401	389	790
Manipulation check (subjective)			
1-2	1	27	28
3-4	23	76	99
5-6	46	82	128
7-8	196	167	363
9-10	179	88	267

Table 2. Distribution of Key Demographic Variables by Group (Counts)

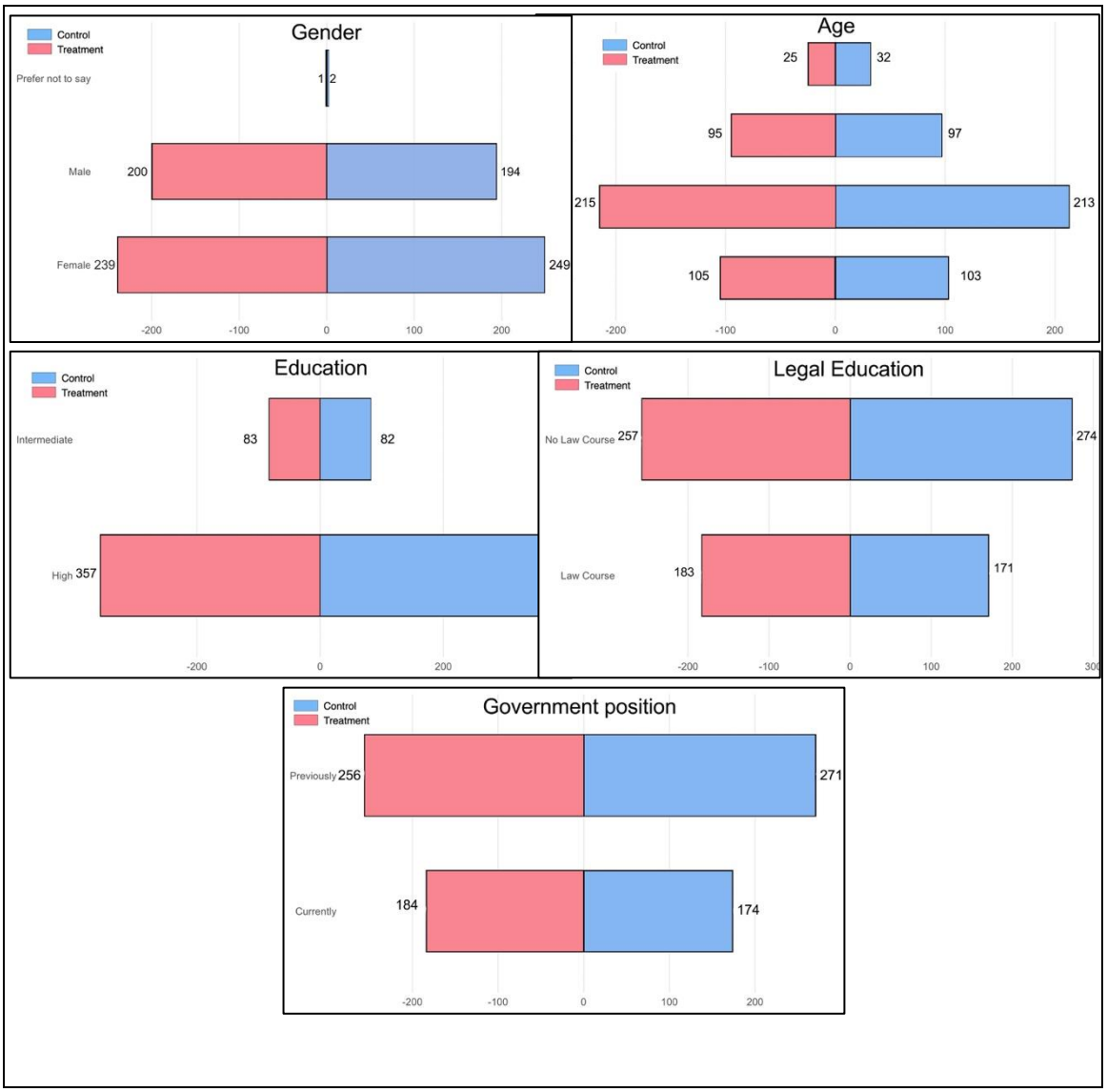


Table 3. Randomization Check

Variable	Control group	Treatment group	Chi² (df)	p-value
Gender			$\chi^2(2) = 0.60$	0.740
Female	56%	54%		
Male	44%	46%		
Prefer not to say	0.4%	0.2%		
Age			$\chi^2(3) = 0.88$	0.830
16–29 years	23%	24%		
30–49 years	48%	49%		
50–64 years	22%	22%		
65–99 years	7%	6%		
Education Level			$\chi^2(1) = 0.03$	0.868
Intermediate	18%	19%		
High	82%	81%		
Legal Education			$\chi^2(1) = 0.92$	0.337
Taken Law Course	38%	42%		
No Law Course	62%	58%		
Government pos.			$\chi^2(1) = 0.68$	0.410
Currently in Government	39%	42%		
Previously in Government	61%	58%		

Note. The distributions are reported in percentages. Differences across groups were assessed using a Chi-square test for homogeneity (Agresti 2018: 225–239). None of the tests yielded statistically significant results, indicating no systematic differences between groups and suggesting that the randomization was successful.

Appendix E. Participants Country of Residence

Table 1. Frequency of Participants per Country by Group (Counts)

Country of residence *n* = 33	**Control**	**Treatment**	**Total**
United Kingdom	175	177	352
South Africa	119	129	248
United States	66	44	110
Poland	17	11	28
Canada	10	10	20
Australia	9	6	15
Germany	3	12	15
Kenya	7	6	13
Italy	7	4	11
Portugal	4	9	13
Hungary	3	4	7
Netherlands	3	4	7
Greece	2	3	5
Spain	3	2	5
France	0	4	4
Ireland	3	1	4
Sweden	3	1	4
Belgium	1	1	2
Brazil	0	2	2
Denmark	2	0	2
Japan	1	1	2
Austria	1	1	2
New Zealand	0	2	2
Slovenia	2	0	2
Czech Republic	0	1	1
Estonia	0	1	1
India	1	0	1
Indonesia	1	0	1
Israel	1	0	1
Latvia	0	1	1
Malaysia	1	0	1
Mexico	0	1	1
Norway	0	1	1
Other	0	1	1