

EUR Research Information Portal

Ideological bias in constitutional judgments

Published in:

Journal of Empirical Legal Studies

Publication status and date:

Published: 01/09/2022

DOI (link to publisher):

[10.1111/jels.12323](https://doi.org/10.1111/jels.12323)

Document Version

Publisher's PDF, also known as Version of record

Document License/Available under:

CC BY

Citation for the published version (APA):

Kantorowicz-Reznichenko, E., Kantorowicz, J., & Weinshall, K. (2022). Ideological bias in constitutional judgments: Experimental analysis and potential solutions. *Journal of Empirical Legal Studies*, 19(3), 716-757. <https://doi.org/10.1111/jels.12323>

[Link to publication on the EUR Research Information Portal](#)

Terms and Conditions of Use

Except as permitted by the applicable copyright law, you may not reproduce or make this material available to any third party without the prior written permission from the copyright holder(s). Copyright law allows the following uses of this material without prior permission:

- you may download, save and print a copy of this material for your personal use only;
- you may share the EUR portal link to this material.

In case the material is published with an open access license (e.g. a Creative Commons (CC) license), other uses may be allowed. Please check the terms and conditions of the specific license.

Take-down policy

If you believe that this material infringes your copyright and/or any other intellectual property rights, you may request its removal by contacting us at the following email address: openaccess.library@eur.nl. Please provide us with all the relevant information, including the reasons why you believe any of your rights have been infringed. In case of a legitimate complaint, we will make the material inaccessible and/or remove it from the website.

Ideological bias in constitutional judgments: Experimental analysis and potential solutions

Elena Kantorowicz-Reznichenko¹ | Jarosław Kantorowicz² | Keren Weinshall³

¹Chair of Quantitative Empirical Legal Studies, Rotterdam Institute of Law and Economics, Erasmus School of Law, Erasmus University Rotterdam, Rotterdam, The Netherlands

²Institute of Security and Global Affairs and Department of Economics, Leiden University, Leiden, The Netherlands

³Edward S. Silver Chair in Civil Procedure, Faculty of Law, The Hebrew University of Jerusalem, Jerusalem, Israel

Correspondence

Elena Kantorowicz-Reznichenko, Chair of Quantitative Empirical Legal Studies, Rotterdam Institute of Law and Economics, Erasmus School of Law, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands.

Email: reznichenko@law.eur.nl

Abstract

Despite the importance and neutrality of constitutional rights, empirical research suggests that ideological inclinations unduly affect their assessment and application. In this study, we conducted two experiments in order to investigate the nature of the ideological bias in a constitutionally relevant decision (right-to-demonstration), and how to mitigate it. We find that ideological bias is driven by in-group favoritism. In addition, we find that prior commitment, through a signed declaration, to be impartial or to prioritize constitutional rights encourages participants not to disfavor out-groups. On the other hand, we do not find evidence that using a temporary blinding procedure mitigates the ideological bias.

INTRODUCTION

The principle of universality of rights is the cornerstone of constitutional and international human rights law.¹ According to this principle, all human beings are entitled to inherent rights. As such, rights equally protect all individuals, regardless of their personal status, attributes, beliefs and ideological or political inclination (Donnelly, 2007). In particular, the question of ideology or political affiliation should not affect the application of political rights. The ability to

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Journal of Empirical Legal Studies* published by Cornell Law School and Wiley Periodicals LLC.

participate in civil and political life is by nature meant to be afforded equally, in a content-neutral manner.

However, no right is unlimited. Proportionality is one of the most common meta-constitutional principles for balancing and adjudicating between human rights that conflict with other public interests (Alexy, 2010, pp. 66–69, Cohen-Eliya & Porat, 2011, Huscroft et al., 2014, Jackson, 2014, Kremnitzer et al., 2020; Sweet & Mathews, 2008, p. 113). Despite the prevalence of basic legal principles as the universality of rights and proportionality, research suggests that ideological inclinations unduly affect the assessment and balancing of constitutional rights. Critical legal scholars claim that the abstractness of the proportionality principle produces inconsistent application of rights (Choudhry, 2006; Kumm & Walen, 2014) often influenced by the policy preference of the judge or decision maker.

Empirical evidence supports these claims. For example, Sulitzeanu-Kenan, Kremnitzer, and Alon find that legal proportionality judgments are strongly influenced by adjudicators' ideological preferences (Sulitzeanu-Kenan et al., 2016). Kahan et al. (2012) show that political affiliation affects people's perceptions of facts relevant to distinguishing constitutionally protected "speech" from unprotected "conduct." Their findings demonstrate the high impact of "cultural cognition," which they define as the unconscious influence of individuals' ideological group on their perceptions of legally consequential facts. Likewise, Epstein et al. (2018) reveal an ideological in-group bias in freedom-of-expression cases decided by US Supreme Court justices between the years 1953 and 2014. Both liberal and conservative justices tend to more vigorously protect the free speech rights of groups that are aligned with their own political inclinations.

The increasing polarization of societies, on many ideologically sensitive issues, exacerbate the problem of ideological bias. Yet, despite the extensive research on ideological bias, its nature in the context of constitutional rights is unclear. Is it primarily driven by in-group love or by out-group hate? One case of discrimination based on ideological bias would be to accord special rights to one's own group—even if other groups still receive a fair, albeit lesser, allocation of rights. Another is to allow a "baseline" of rights to all groups, but negatively discriminate against (some) out-groups. Furthermore, while the problem has been identified—the influence of ideological bias in such decisions—its potential solutions have not been investigated.

This paper aims to examine the nature of ideological bias in rights-based judgments and to consider potential solutions. For this purpose, we conducted two survey experiments in the context of freedom of speech, freedom of assembly and the right to demonstrate. We chose to conduct the experiments in Poland, a setting with rising political polarization, where the role of constitutional rights is being challenged.² In such an environment, ideological bias can play a significant role in the decisions of public officials and its mitigation is critically important.

In both experiments, we manipulated the ideological distance of participants vis-à-vis an organization applying for a permit to demonstrate, as well as the decision-making procedure that leads them to approve or reject the application. We employ two different behavioral interventions intended to “debias” the ideological bias through legal rules (Jolls & Sunstein, 2006; Rachlinski & Jourden, 1998; Thaler & Sunstein, 2008).

In the first experiment, the identity of the organization seeking approval to demonstrate was revealed to the treatment group only after they expressed their general perception of the importance of the right to demonstrate. This temporarily blinding procedure was expected to mitigate ideological bias by getting participants to make an affirmation that commits them to their initial unbiased preference and exploiting the natural desire to avoid cognitive dissonance. The ideological distance spectrum we created also allowed us to examine the nature of the ideological bias in this study. By introducing a treatment group that was not initially exposed to the identity of the organization, we were able to create a neutral and unbiased group. Then we compared the decisions of the unbiased group to the ideologically “affected” group. This comparison allowed us to identify whether the ideological bias was driven by in-group love, out-group hate, or both.

In the second experiment, we randomly allocated participants to read, or read and sign, a declaration prior to making a decision. The declaration stated that their decision would be impartial or that they would give appropriate weight to constitutional rights. This behavioral intervention was aimed at stimulating a self-monitoring mechanism by making the “morality” of the decision more salient. Given the saliency of this moral commitment, participants in the treatments groups were expected to give less weight to their own ideological view.

Findings from both experiments show that respondents who are ideologically proximate to the applying organization are much more likely to approve their request to demonstrate relative to respondents who are ideologically distant. This was consistent across the two experiments and confirms the existence of the ideological bias. Results from the first experiment (481 participants) suggest that the ideological bias mechanism is driven by in-group love and not by out-group hate. At the same time, we did not find evidence that the behavioral intervention we attempted in the first experiment is effective in mitigating the ideological bias. Before knowing who has applied for the permit to protest, the subjects’ stance on the right to free speech (demonstration) was independent from their ideological stance. Afterward, however, participants in both the treatment and control group demonstrated equally strong ideological bias. The additional finding that the debiasing procedure did not mitigate the ideological bias stresses the prevalence of such bias. On the other hand, the behavioral intervention employed in the second experiment (1952 participants) was able to mitigate to a certain extent the ideological bias for participants who read and signed impartiality or constitutional rights declarations, or even those who just read

the declarations. However, even with this intervention, we did not succeed in eliminating the ideological bias of the most aligned participants to the applying organization.

The study contributes to the existing literature in several ways. First, it demonstrates the robust influence of ideology on decisions about basic rights. The no-effect results in Experiment 1, and the inability to overcome the ideological bias toward groups who are the most ideologically aligned with the participants in Experiment 2, can further stress the potency of ideology, especially considering the sample size and its statistical power. Second, we expand the theoretical understanding of the nature of ideological bias in the context of constitutional rights. To the best of our knowledge, our experiment is the first to explicitly test in-group love versus out-group hate in this context, using the behavior of an unbiased group as a benchmark. Other studies have shown the bias of favoring an in-group and disfavoring an out-group. However, without a comparison to a neutral group, the exact direction of the discrimination remains unclear. Third, we offer a more practical contribution with the behavioral intervention developed in our second experiment. We further discuss the policy implications, offering a cheap and easy way to adapt an intervention that can mitigate ideological bias in decision-making processes regarding the application of constitutional rights. However, before adopting policy reforms, we suggest conducting additional research to enhance external validity, chiefly in expanding the studied countries, context and type of participants.

The paper is structured as follows. Section 2 outlines the legal framework in general, and in Poland in particular, which is the context of our experiments. In Section 3, we provide the theoretical framework and discuss the empirical literature on ideological biases and motivated reasoning in legal judgment, as well as describe the state of the art on in-group–out-group bias, cognitive dissonance, and self-awareness theory. Based on the conceptual framework, we present our behavioral predictions. In Sections 4 and 5, we describe the experimental design, data and results of our two experiments. Finally, Section 6 concludes with a discussion of policy implications, the limitations of our study and ideas for future research.

CONSTITUTIONAL RIGHTS AND THE PRINCIPLE OF PROPORTIONALITY IN POLAND

The principle of proportionality, which originated in 19th-century German jurisprudence, migrated as a central constitutional institution throughout Europe (Pulido, 2013) and was later adopted in countries across Asia,³ North and Latin America,⁴ Africa,⁵ and Australia (Carter, 2016).⁶ Proportionality also became a central principle of international law, as expressed for example in the European



Convention on Human Rights and in World Trade Organization regulations (Andenas & Zleptnig, 2007; Harbo, 2010; Kumm, 2007).

The principle is intended to guide policy makers (and then judges) in evaluating policy alternatives by structuring their decision in a series of stages. These characteristically include four steps⁷: confirming the existence of a legitimate or worthy public purpose for limiting rights; employing the suitability test to check a rational connection between the limiting means and the public purpose; using the necessity test to examine whether the public purpose can be attained by a less restrictive alternative; and invoking the final balancing test to weigh the benefit of the public policy relative to the costs incurred by infringement of the right (Kremnitzer et al., 2020).

Although the principle of proportionality—understood as the prohibition on undue interference by public authorities—was well known in Polish administrative and criminal law, its constitutional status was recognized in the rulings of the Constitutional Tribunal only in the early 1990s (particularly rulings P 1/87, K 6/90 and K 10/92, see in Śledzińska-Simon, 2020). After the promulgation of the new Polish Constitution in 1997, two distinct applications of the proportionality principle were developed: one with regard to excessive interference and limitations imposed by authorities on other public bodies (for instance, municipalities), and the other regarding restrictions on constitutional rights based on the general limitation clause. The latter application is of interest given the context of this study—the right of demonstration and peaceful assembly. It is important to note that while the proportionality principle serves as an assessment criterion in judicial practice and as a method of argumentation to justify a ruling in a given case, this rule is less formalized in the sphere of public administration, where it is expressed in general guidelines for issuing laws and regulations that are the least burdensome for society (Śledzińska-Simon, 2019, pp. 72–74).

In line with most authoritative international human rights treaties (such as the European Convention of Human Rights), the Polish Constitution of 1997 guarantees the right to demonstrate. Article 57 guarantees the freedom of peaceful assembly to everyone. On the other hand, and this is where the principle of proportionality comes into play, the Constitution's limitation clause allows for restrictions of these rights. As stipulated in article 31 paragraph 3, any limitation on the exercise of constitutional freedoms and rights may only be imposed by statute, and only when necessary in a democratic state for *the protection of its security or public order, or to protect the natural environment, health or public morals, or the freedoms and rights of other persons*. Such limitations must not violate the essence of freedoms and rights.

Detailed rules and procedures regarding public gatherings are stipulated in the 2015 Public Gatherings (Assemblies) Law. Article 7.1 of the law states that an organizer of a public gathering must inform the municipal authority (essentially, the mayor) no earlier than 30 days and no later than 6 days before the

gathering is slated to take place. Since 2002, mayors in Poland have been directly elected. In large cities, they may be members of national political parties (as in Warsaw, for example). Thus, the issue of organizing gatherings may be heavily politicized. The law further specifies in article 14 that a municipal authority may decide (*ex-ante*) to ban a public gathering no later than 96 h before the gathering takes place, if: (1) its aim interferes with the freedom of a peaceful gathering, or its rules and aims interfere with criminal regulation; (2) the gathering may pose a threat to life or public health, or may result in substantial material loss; (3) the gathering is slated to occur at the same time and place as a cyclical gathering.⁸ It is not uncommon for municipal authorities to resort to article 14 to ban gatherings or attempt to ban them.

Case law includes several decisions by municipal authorities and the confirmation of these decisions by heads of *voivodships* (a level of government above the municipal level) which were overturned by administrative and ordinary courts.

One example is Case I OSK 2538/13, in which the Supreme Administrative Court overturned the municipal authority's decision to ban a public gathering of around 20 people in the city center of Kielce. The gathering was intended to protest the use of city funds to support pro-abortion organizations and promote antiabortion messages. The municipal authority invoked article 14 (2) stating that the gathering, organized in a very busy part of the city, may impose difficulties on pedestrians and interfere with car traffic and general public order. Thus, it may pose a threat to the life and health of people and may result in significant material losses. As "evidence," the police presented statistics showing the high frequency of traffic accidents in that part of the city. The Supreme Administrative Court stated in its ruling that the freedom of assembly is an essential value in democratic society; it is a right granted by the Polish Constitution and by international law ratified by Poland.

In another case (Case K 21/05), the Constitutional Tribunal asserted that public authorities are obliged to guarantee the security of anyone lawfully exerting their right to public assembly. While municipal authorities may restrict this right if they identify threats to life and public, they cannot just formulate suppositions regarding potential threats. They need to conduct a robust analysis referring to a concrete case and circumstances and must show that the threat to life and health in a given case is real. The Constitutional Tribunal ruled in this case that the municipal authorities had failed to comply with this norm. It is noteworthy that the gathering discussed in this case had a static character and only 20 people were supposed to be present. From this perspective, the assessment of potential threat did not correspond with the circumstances of this case. According to the reasoning presented by the court, if any interference with car traffic could be used to justify the restriction of freedom of assembly, no gathering could be effectively organized. The court further stipulated that when assessing restrictions to freedom of assembly, this constitutional right should

take precedence over the rights and freedoms of others (e.g., difficulties in getting home or in gaining access to banks and other services).⁹

THEORETICAL FRAMEWORK: BIASES THAT CAUSE, AND THAT MIGHT SOLVE THE PROBLEM

The nature of ideological biases and motivated reasoning in legal judgments

Previous studies found strong evidence for ideological and political bias in the application of political rights. Several overlapping mechanisms have been discussed when explaining this bias. A primary explanation is the effect of intergroup bias, sometimes known as in-group favoritism. A systematic tendency to evaluate one's own group and its members more favorably than a non-membership group (an out-group) and its members was found in various social grouping (Hewstone et al., 2002; Tajfel, 1981). This includes trivial alliances, for example, groups based on a preference for paintings by Klee or Kandinsky (Tajfel et al., 1971), as well as groups defined by race, religiosity, ethnicity, and so on (Allport et al., 1954, pp. 29–36). Intergroup bias is especially salient in morality-based groups, defined by their shared beliefs involving judgments of harm (Parker & Janoff-Bulman, 2013; Weisel & Böhm, 2015). Various studies have shown that liberals positively evaluate liberals (their in-group) and hold negative attitudes toward conservatives (their out-group), and vice versa (Huber & Malhotra, 2017; Rogowski & Sutherland, 2016). Accordingly, liberals are more generous toward those who share their ideological commitments (Rand et al., 2009), and more willing to accept conservative policies if they are proposed by members of their in-group (Cohen, 2003).

When making judgments that may impact one's ideological interests, an ideological bias is also influenced by motivated reasoning, which refers to the selective processing and evaluation of evidence with the desired solution in mind. Motivated political or moral reasoning is usually a post hoc construction, generated after a judgment has been reached (Haidt, 2001). People tend to evaluate information through a partisan-ideological lens, discounting evidence presented by a political rival and positively evaluating the same content when it is offered by a co-partisan (Bolsen et al., 2014; Goren et al., 2009; Leeper & Slothuus, 2014). They are inclined to seek confirmatory evidence that is congruent with their own ideological views (Taber & Lodge, 2006).

Motivated reasoning accentuates the importance of moral, ideological, or political in-group influences. People endorse whichever position reinforces their connection to other members in a group whose beliefs are congenial to their own worldviews. They interpret new evidence in a biased way that reinforces their predispositions. As a result, groups with opposing values often become

more polarized when exposed to information such as scientifically sound data about climate change (Kahan, 2010; Kraft et al., 2015).

Much like the case of polarized beliefs/skepticism toward scientific evidence, previous research has demonstrated the effects of intergroup bias on the assessment of legally consequential facts, arguments, and decisions.¹⁰ Political and ideological bias in particular was found in different legal settings. For example, in the criminal law setting, Sood and Darley (2012) showed that both liberals and conservatives are more apt to cite the legal principle of harm (which suggests that the state should criminalize conduct only when necessary to prevent harm) to criminalize and severely punish an out-group member as compared to an in-group member.

In the constitutional law arena, Furgeson et al. (2008) asked law students to determine the constitutionality of a hypothetical law; the researchers manipulated the policy implications of the law, but kept all legal evidence constant. They found that liberal law students were more likely to overturn laws that lowered taxes in comparison to laws that raised taxes, and vice versa for conservatives. The intergroup effect remained significant even after employing an incentive to select the ruling best supported by the legal evidence. Likewise, Cope and Crabtree (2020) found that attitudes toward the implementation of international human rights were driven by the political-party affiliation. Sulitzeanu-Kenan et al. (2016) show that ideological preferences strongly influence legal expert's judgments regarding the proportionality of hypothetical scenarios describing targeted killings of terrorists (Sulitzeanu-Kenan et al., 2016).

Furthermore, several studies have found ideological bias in constitutional assessments related to political rights. Kahan et al. (2012) displayed a video of a political demonstration and told half the participants that the demonstrators were protesting abortion outside an abortion clinic, and the other half that the demonstrators were protesting the military's "don't ask, don't tell" policy outside a military center. Conservatives identified the anti-military (out-group) protest as violent and the anti-abortion (in-group) protest as peaceful, while liberals had the opposite perception. Similarly, in an ongoing experimental study, Sulitzeanu-Kenan et al. (2022) found that in both Canada and Israel leftist respondents view left-wing protest rallies more favorably than right-wing protest rallies, while the opposite is true among right-wing respondents. Finally, using observational methods, a recent study by Epstein et al. (2018) found that both liberal and conservative US Supreme Court justices tend to more vigorously protect the free speech rights of groups that are aligned with their own political inclinations.

Despite the consensus in the literature on the existence of the ideological bias, including in legal judgments, it is not yet clear what is the main direction of this bias in constitutional judgments. In other words, is it the in-group favoritism, or the out-group derogation, that mainly drives the decision of the individual?



The general literature on intergroup conflict suggests that the strongest channel for intergroup bias is in-group love, rather than out-group derogation (Halevy et al., 2008, 2012; Mummendey & Otten, 1998; Weisel & Böhm, 2015; Yamagishi & Mifune, 2009). This was suggested by Allport in his seminal book *The Nature of Prejudice*, which attributed primary importance to in-group formation. An in-group is formed out of a sense of loyalty; this can be to a person's family, religious group, political affiliation, and so on (Allport et al., 1954, pp. 29–36). The out-group's existence can strengthen the in-group's sense of belonging and superiority, but it is not necessarily accompanied by out-group hostility (Allport et al., 1954, pp. 41–42, 47, Brewer, 1999). A recent meta-analysis of intergroup discrimination confirmed that discrimination is driven by in-group love rather than out-group hate or derogation (Balliet et al., 2014). Nevertheless, another study suggested that morality-based groups (polarized political affiliation) are a special case and might lead to different results than those found in general intergroup conflict (Parker & Janoff-Bulman, 2013, pp. 82–84). This was later confirmed in a series of experiments. The authors found that in the context of morality-based groups, group bias was equally driven by in-group love and out-group hate. Furthermore, when out-group hate was expressed by help-avoidance as opposed to active harm, morality-based groups expressed stronger out-group hate (Weisel & Böhm, 2015).

Given the supporting evidence in the literature discussed above, we first test whether ideological bias exists in the context of permission to demonstrate. Since the existence of the ideological bias in constitutional questions is central to this paper, this hypothesis was also tested in Experiment 2.

Hypothesis 1 (H1). *The weight participants attribute to constitutional rights, and their willingness to approve a political demonstration, will increase for ideological groups that are proximate to the participants' ideological views (their in-group) and decrease for ideologically distant groups.*

From the empirical literature it is still unclear, however, whether ideological bias in legal and political decisions is driven solely, or mainly, by in-group love or by out-group hate. In other words, the question remains: Do decision makers in the context of constitutional rights simply favor their own ideological groups and treat other groups neutrally, or do they also treat other groups unfavorably?

In our two experiments, we used morality-based groups. In light of the exceptional behavior of these groups, as noted above, we formulate the following hypothesis:

Hypothesis 2 (H2). *Participants will give more weight to the constitutional rights of their ideological in-group as compared to a neutral*

assessment when the identity of the applicant is unknown. In addition, participants will give less weight to the constitutional rights of an ideological out-group as compared to a neutral assessment when the identity of the applicant is unknown. [a combination of in-group love and out-group hate].

Cognitive biases which can be used for “debiasing”

In this paper, aside from investigating the nature of the ideological bias, we also examine potential ways to mitigate it. In particular, we look at different behavioral interventions in order to try to debias decision makers and mitigate the impact of their ideological bias. In this section, we discuss the theoretical underpinnings of those behavioral interventions.

Blinding procedures

The idea behind blinding procedures is to conceal from the decision maker certain information that is irrelevant for the decision, but nonetheless often affects it. Such procedures have been used and tested in different contexts. For instance, gender bias was observed in the process of choosing musicians for the New York Philharmonic Orchestra. To overcome this bias, the examiners started conducting auditions without seeing the candidate. As a result, the number of women accepted into the orchestra increased (Goldin & Rouse, 2000). Another example is the double-blind procedure used in medical clinical trials: Neither the experimenter nor the participants know who received the real treatment and who received a placebo. Blinding procedures are designed to reduce the investigators' bias in the assessment of data and participants' bias in the psychological response to the intervention (Schulz & Grimes, 2002).

Blinding procedures have also been discussed in the context of the criminal justice system. Racial bias, often unconscious, has been observed in the American criminal justice system, where black prisoners are overrepresented in the prison population relative to their proportion in the general population. One study proposed a blinding procedure in which prosecutors decide whether to file criminal charges without knowing the race of the suspect (Sah et al., 2015).

In the context of litigation, a blinding procedure designed to overcome a bias toward self-interest was tested and confirmed that people assess future trial outcomes and fairness from the prism of their own interest. Thus, the same set of information is assessed differently by the plaintiff and the defendant. This makes it difficult to settle cases and leads to a proliferation of litigation cases. In one study, participants received information about a hypothetical legal case and were asked to evaluate the expected trial outcome and to negotiate a



settlement—without knowing whether they were playing the role of the plaintiff or the defendant. These “blinded” participants managed to reach a settlement more frequently than participants who knew the role they were playing (Babcock et al., 1997).

Blinding procedures seem to constitute an efficient strategy for eliminating bias, but it was not possible to employ full blinding in the context of our study on approval of public demonstrations. The identity of the applying organization is less relevant to the question of the constitutional right (demonstration) but is required to assess the question of possible threat to public order. Some organizations evoke more public disturbance than others, as experience shows, and as other unique cultural factors indicate. Therefore, it is essential to know the identity of the applying organization before deciding whether to approve its request to demonstrate. We thus use a “temporarily blinding procedure,” or more precisely a procedure where we postpone the exposure of participants to the identity of the group.¹¹ This process allowed us to utilize the phenomenon of cognitive dissonance as a debiasing strategy.

Cognitive dissonance

According to the theory of cognitive dissonance, people have a strong inclination for consistency between their attitudes or opinions and their behavior, and the psychological feeling of dissonance arises when there is inconsistency between those “pieces of knowledge.” Consequently, people try to reduce this inconsistency in order to alleviate the sense of dissonance. There are two ways to reduce such dissonance: (1) to change one’s attitudes to fit one’s behavior, or (2) to change one’s behavior to fit one’s attitudes. For example, smokers who know that smoking is bad for their health can convince themselves that smoking is not so harmful after all, or they can stop smoking. In both scenarios, the dissonance is reduced (Festinger, 1962, pp. 1–6).

Early studies demonstrated that cognitive dissonance can be so powerful that forcing compliance with something that contradicts a person’s attitudes can induce that person to change his or her attitudes. For example, writing an essay supporting an opposing view in exchange for a small incentive later resulted in stronger support for that view. In another study, subjects who participated in a boring task and then agreed to tell other participants that the task was actually fun ultimately viewed the task more favorably (Festinger & Carlsmith, 1959, Brehm & Cohen, 1962, pp. 73–78). Similar insights were found in experiments measuring how suffering or a threat relates to cognitive dissonance. For instance, when people suffer in order to participate in an unworthy activity, they experience a sense of dissonance, asking themselves: Why should I put myself through such suffering for such an unworthy activity? In order to reduce this

dissonance and justify their effort, people simply change the way they perceive the activity, and start seeing it as worthier (Cooper & Hogg, 2007, pp. 20–26).

Recent studies have provided further support for cognitive dissonance theory. For example, one study used a dictator game design¹² to examine whether unfair allocation of collective earnings would be followed by self-deception to reduce the dissonance between wanting more money and the desire to act fairly. This study indeed demonstrated that many participants who chose to allocate the money unfairly and tried to justify their decision through self-deception, made a similar decision even when they did not have a personal stake in the allocation. This implies that after behaving unfairly, they changed their perception of what is fair and later adjusted their behavior to be consistent with their changed perception—even when there was no material incentive to do so (Konow, 2000).

In light of the reviewed theoretical framework, we made the following hypothesis:

Hypothesis 3 (H3). *A temporary blinding procedure, which utilizes cognitive dissonance, will reduce participants' ideological bias when deciding whether to authorize a demonstration.*

Objective self-awareness theory

An additional possibility to mitigate ideological bias might be making the moral or professional obligation more salient. One relevant theory from psychology for this context is *objective self-awareness theory*. Simply stated, this theory refers to situations in which people focus attention on their “self” and assess the correctness of their behavior by automatically comparing it with standard behavior—what is correct to do. In other words, people compare their ideal self to their actual self. Inconsistency between the self and the standard generates discomfort. This discomfort can be reduced by either changing one’s current behavior to fit the standard (which is relatively fixed and hard to change), or by diverting one’s attention from the stimuli that trigger the comparison (Duval & Wicklund, 1972, pp. 3–4, 10).

Individual self-monitoring behavior under the threat of external review is exhibited in many situations—for instance, when filing tax returns or insurance claims. The incentives to misreport are especially strong when the probability of review is low. The decisions of public officials also rely on a self-monitoring system in lieu of constitutional review. In particular, public officials are expected to make impartial decisions and not to be driven by their ideological views. Especially when balancing between constitutional rights and public interests, public officials are expected to give primacy to the former and not allow irrelevant factors to influence their decision. Relying on the objective self-awareness theory,



making their obligation salient just before the public officials need to make their decision has the potential to reduce their bias. The obligation would serve as the standard to which the decision makers will potentially adjust their behavior.¹³ In light of the behavioral insights, we formulated the following hypothesis:

Hypothesis 4 (H4). *The likelihood of approving a demonstration will depend less on the ideological distance between the decision makers (participants) and the applying organization when participants are asked to read and sign a declaration, or just read a declaration, before making their decision, as compared to when they are not requested to do so.*

EXPERIMENT 1: THE NATURE OF THE IDEOLOGICAL BIAS, TEMPORARY BLINDING AND COGNITIVE DISSONANCE

Experiment 1 served to examine hypotheses H₁–H₃. In particular, we have designed our experiment to first investigate whether ideological bias exists in the context of constitutional rights (H₁). If yes, we were then interested to examine what is the nature of this bias, and whether it is driven by in-group love, out-group hate, or both (H₂). Finally, our experiment was aimed to test whether using temporary blinding procedure and utilizing the phenomenon of cognitive dissonance, would mitigate the ideological bias (H₃).

Experimental design

Each participant was asked to imagine that he or she were a public official (mayor) in a large Polish city and had to decide whether to approve an application to hold a demonstration.¹⁴ One applying organization had a conservative (anti-abortion) ideology and the other a liberal (pro-choice) ideology.¹⁵ After the main questions in the study, we also assessed the participants' ideological stance regarding abortion on a scale from 0 to 10 where 0 = liberal and 10 = conservative.¹⁶ This enabled us to measure each participant's ideological distance from the applying organizations, whose ideological positions were assumed to be at the extremes—that is, 0 for the pro-choice (liberal) organization and 10 for the anti-abortion (conservative) organization. The resulting “ideological distance” variable is thus measured on a scale of 0–10, with 0 indicating a complete ideological match between a respondent and the applying organization (ideologically aligned) and 10 expressing a complete ideological mismatch (ideologically misaligned). Accordingly, the scores between 0 and 10 represent various degrees of (mis)alignment.

We were interested in examining: (1) the nature of the ideological bias, and (2) a potential method for mitigating the bias. For this purpose, we randomly divided the participants into four groups—control group/conservative applicant; control group/liberal applicant; treatment group/conservative applicant; and treatment group/liberal applicant.

All participants were presented with two main questions and a short questionnaire. The first question asked participants to rate the importance of the right to demonstrate (Q1)¹⁷ and the second was about the actual decision on the demonstration application (Q2).¹⁸ In Figure 1, we present the flow of the experiment for all groups. The full text of the experiment can be found in Table E1.1 in the Supporting Information.

The purpose of Q1 was twofold. First, it allowed us to examine the nature of the ideological bias. The treatment group was neutral in the sense that they were presented with the question without the background of a particular organization, and therefore they were not driven by a particular agenda in answering the question. Thus, their answer presented a benchmark evaluation, representing a non-biased opinion. We then could compare it to the “ideologically exposed” control group and identify the features of the bias (H₂). If the control group, in comparison to the treatment group, attributed greater importance to the right to demonstrate in the case of an ideologically aligned applying organization, this would suggest in-group love. Conversely, if the control group attributed less

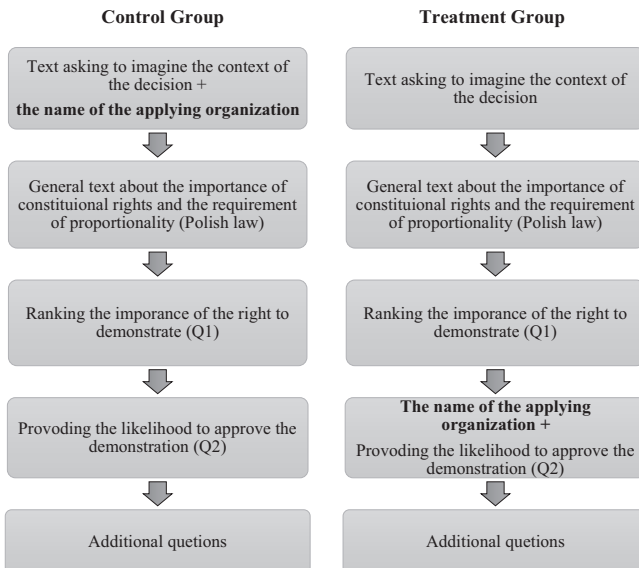


FIGURE 1 Experimental flow

importance to the right to demonstrate in the case of an ideologically misaligned applying organization, this would suggest out-group hate.

Second, this question was designed to commit the participants in the treatment group to a more neutral mode of decision-making. Applying cognitive dissonance theory, we predicted that answering the first question in a neutral manner might mitigate bias in the participants' responses to the second question, where they had to decide whether to approve the demonstration application. In other words, we expected that the level of importance a person ascribed to the right to demonstrate would influence his or her decision on demonstration approval (Q2). The need to avoid cognitive dissonance might lead the participant to align his or her responses to Q1 and Q2. For example, cognitive dissonance would arise if a person initially stated that the right to demonstrate should prevail, but then expressed reluctance to authorize a demonstration by an ideologically distant organization. That is, the person's general belief in the importance of constitutional rights would be dissonant with his or her decision in the particular case.¹⁹

We expected to find support for our hypothesis of bias mitigation (H₃) if the treatment group, in comparison to the control group, expressed a greater inclination to approve demonstrations in cases of ideological misalignment and/or a lower likelihood of demonstration approval in cases of ideological alignment.

Experimental procedure

We used a representative sample (in terms of age, gender and education) of the Polish population, which was recruited by a professional survey firm, Kantar Polska. The sample included 481 participants ($N = 481$), of which 49% were female, 33% highly educated (bachelor's or master's degree) and 37% nonurban. The average age recorded in the sample was nearly 44 and the average position on the 0–10 ideological scale regarding abortion was 3.53 (where 0 = pro-choice and 10 = antiabortion). The study was conducted on December 17–19, 2018. The balancing tests and detailed descriptive statistics are provided in Tables E1.2 and E1.3 in the Supporting Information.

Figure 2 shows the frequency distribution of the ideological distance variable, which is generated as an absolute distance between the respondents' ideological position on the abortion question and the ideological position of the applying organization. This variable constitutes a crucial component of the empirical model. The distribution has three modes. Roughly 20% of respondents are perfectly aligned with the applying organization (a score of 0 on the ideological difference scale); nearly 12% of respondents are in the middle (a score of 5) and 19% are perfectly misaligned (a score of 10). Approximately, the same number of people is located above and below the middle position (43% and 45%, respectively).²⁰

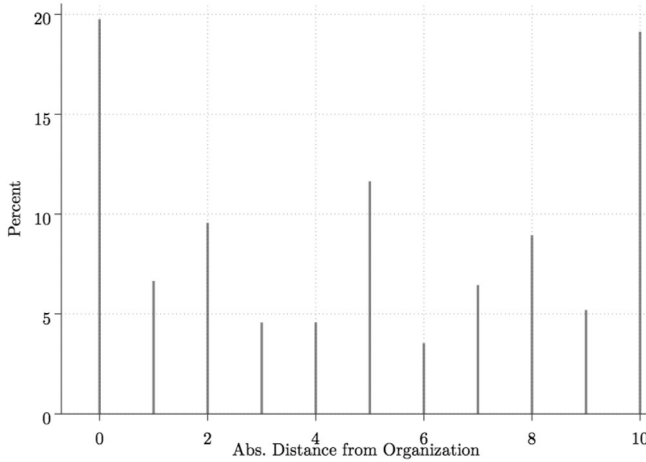


FIGURE 2 Distribution of the ideological distance variable in Experiment 1

Results

To test for hypotheses H₁–H₃, we estimated a range of ordinary least squares (OLS) regressions. The interaction effect model specified in Equation 1 displayed below boils down to estimating two regression lines showing the relationship between the importance of the constitutional right to demonstrate or the likelihood of approving a demonstration, on the one hand, and the ideological distance of respondents, on the other hand, for both the control and treatment groups. The model is specified as follows:

$$Y_i = \alpha_0 + \beta_1 D_i + \beta_2 Dist_i + \delta_1 D_i \times Dist_i + \varepsilon_i, \quad (1)$$

where Y_i is a dependent (outcome) variable—that is, the importance of the right to demonstrate or the likelihood of approving the demonstration. D_i is our treatment variable, which takes two values, 1 for temporary blind decision-making process and 0 for the control group. $Dist_i$ represents the distance between the applicant organization and respondents' position on the abortion ideological question. Lastly, ε_i is an error term. In line with H₁, we expected a negative relationship between the importance assigned to the constitutional right to demonstrate and ideological distance in the control group (when the identity of the organizing entity was known). In other words, in the control group, we expected that the weight respondents assigned to this particular constitutional right would depend on the ideological distance from the applying organization. Consequently, β_2 was expected to be negative. For the treatment group (the

blinded group), we thought that this negative relationship would not hold. That is, since the respondents would not know the organization's identity, they would not assign biased weight to the importance of the constitutional right to demonstrate. As a result, the coefficient next to the interaction term, δ_1 , was anticipated to be positive.

As for H_2 , we expected the ideologically aligned respondents in the control group (who knew the identity of the applying organization) to assign higher weight to the constitutional right, in comparison to their ideological counterparts in the treatment group (who were still unaware of the organization's identity). This corresponds to the in-group love argument and, statistically speaking, for the β_1 to be negative. Likewise, we expected the ideologically misaligned respondents in the control group to assign lower weight to the constitutional right to demonstrate in comparison to their ideological counterparts in the treatment group. This pattern would confirm the existence of out-group hate. This result is not directly visible via the regression coefficients and has to be approached by estimating marginal effects.

Lastly, in line with H_3 , and given the behavioral intervention applied, we expected the aligned respondents from the treatment group to be less likely to authorize the demonstration as compared to the respondents from the control group, and/or the misaligned respondents from the treatment group to be more likely to approve the demonstration as compared to their counterparts in the control group. Statistically speaking, for the aligned respondents this would mean that the β_1 coefficient is negative. For the misaligned respondents, the effect can be assessed by estimating marginal effects.

We start discussing the results by inspecting the visual representation of patterns in Figure 3. In line with H_1 , we see a negative relationship between the perceived importance of the constitutional right to demonstrate and the ideological distance in the control group. The respondents who are ideologically distant from the applying organization assigned less weight to the constitutional right to demonstrate. As displayed in column 1 of Table 1, the coefficient next to the ideological difference variable is negative and the estimate is statistically significant at a level of 5%. For every one-unit shift on the ideological distance scale, the importance assigned to the constitutional right to demonstrate decreases by nearly 0.06. All these results are robust to including various sets of control variables (see columns 2–4 of Table 1).²¹

We can furthermore observe that ideologically aligned respondents in the control group weighed the importance of the right to demonstrate higher than those in the treatment group. Thus, when aligned respondents knew the identity of the applying organization, they were more in favor of the right to demonstrate than when the applying organization was unknown to the participants. This regularity confirms the existence of in-group love, in line with part of H_2 . As seen in column 1 of Table 1, for individuals who are perfectly aligned with the applying organization (a score of 0 on the ideological distance spectrum) the

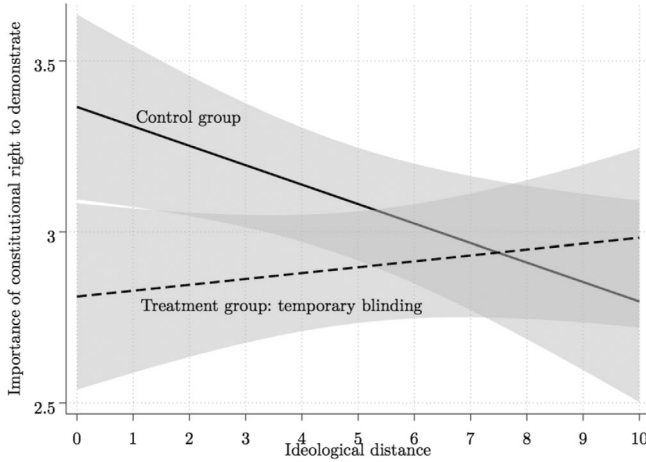


FIGURE 3 Importance of the constitutional right to demonstrate and ideological distance in Experiment 1. The solid line captures the control group, while the dashed line refers to the treatment group. Gray ribbons display 95% confidence intervals.

difference between control and treatment groups is statistically significant at level of 1%. On average, perfectly aligned respondents in the treatment group (unaware of the organization's identity) gave a score that was roughly 0.5 lower than their counterparts in the control group (aware of the organization's identity). Figure 3 indicates that the ideologically misaligned respondents in the control group did not assign lower scores for the importance of the right to demonstrate as compared to those in the treatment group. This suggests we do not find evidence of out-group hate. Figure 3 also shows statistically significant differences between the treatment and control groups up to the score of 4 on the ideological difference scale. The differences between the experimental groups are not statistically significant beyond that point.²² Taken together, the results provide evidence for in-group love and do not support the out-group hate explanation. In other words, the findings suggest that what drives ideological bias in the context of constitutional rights is favoritism toward one's own group, rather than derogation of the other group.

The negative correlation between ideological distance and perceived importance of the right to demonstrate in the control group, and the lack of this correlation in the treatment group serve as a kind of manipulation check of our attempt to harness cognitive dissonance in our procedure. While the participants in the control group were influenced by the identity of the applying organization, participants in the treatment group did not observe this influence and could express their stance on constitutional rights neutrally. This neutral

TABLE 1 Regression model: Importance of the constitutional right to demonstrate

Variables	(1) Baseline	(2) Including controls	(3) Including controls	(4) Full set of controls
Absolute distance from applying organization	-0.057** (0.023)	-0.056** (0.023)	-0.046** (0.022)	-0.046** (0.022)
Treatment: temporal blind	-0.554*** (0.195)	-0.531*** (0.194)	-0.516*** (0.187)	-0.471** (0.191)
Treatment: temporal blind * Abs. distance from organization	0.074** (0.032)	0.068** (0.032)	0.068** (0.030)	0.063** (0.031)
Female		-0.137 (0.119)	-0.087 (0.115)	-0.075 (0.117)
Age		0.006 (0.004)	-0.001 (0.004)	0.001 (0.004)
Highly educated		0.256** (0.128)	0.131 (0.125)	0.183 (0.127)
Nonurban area		0.036 (0.122)	0.125 (0.118)	0.155 (0.120)
General ideology (liberal-conservative)			-0.017 (0.022)	-0.010 (0.023)
General importance of the right to demonstrate			0.117*** (0.026)	0.122*** (0.027)
Active use of the right to demonstrate			0.385*** (0.124)	0.345*** (0.125)
Intense negative attitude on abortion			-0.285** (0.137)	-0.234* (0.140)
Catholic				-0.258* (0.143)
Constant	3.366*** (0.137)	3.050*** (0.243)	2.414*** (0.312)	2.436*** (0.321)
Observations	481	481	481	466
R^2	0.020	0.040	0.122	0.130

Note: SEs in parentheses. Incremental *F*-test provides evidence that the model with an interaction term is better than the model without an interaction term ($F = 5.44$, p -value = 0.020). Column 1 presents baseline regression where we only include dummy variables for the experimental groups, in addition to ideological distance. Column 2 extends the model by controlling for basic individual level characteristics. Column 3 further extends the model and controls for attitudinal positions for “General ideology (liberal-conservative),” “General importance of the right to demonstrate,” “Active use of the right to demonstrate,” “Intense negative attitude on abortion.” For a description of the variables, see Table E1.3 in the Supporting Information.

** $p < 0.1$; * $p < 0.05$; *** $p < 0.01$.

formation of an opinion was meant to anchor the participants before they made the actual decision on whether to approve a specific demonstration.

The second stage of the experiment focused on our main question of interest—the probability of approving a demonstration by a specific organization. We examined if the patterns found for the importance of the right to demonstrate can also be observed when respondents make decisions on whether to authorize a particular demonstration. In other words, we investigated whether the behavioral intervention was successful in mitigating ideological bias. As seen in Figure 4, the lines, which show the relationship between ideological distance and probability of approving the demonstration, for both the control and treatment groups almost overlap, indicating that the patterns are not the same as for the question about the importance of the right to demonstrate. We did not find support that the temporary blinding procedure could achieve its goal of generating cognitive dissonance in order to change the behavioral pattern. In the end, both groups displayed the same trend of approving the demonstration, with a declining likelihood of authorizing the demonstration as the ideological distance widened. Participants in the treatment group seem not to have adjusted their likelihood of approval to their rating of the importance of constitutional rights after learning the identity of the applying organization.

Although the visual illustration of the pattern is rather convincing, in Table 2 we provide an analytical result from the OLS regression model. Whereas the coefficient next to the ideological distance variable is negative and statistically significant at a level of 1%, the interaction term is close to zero and

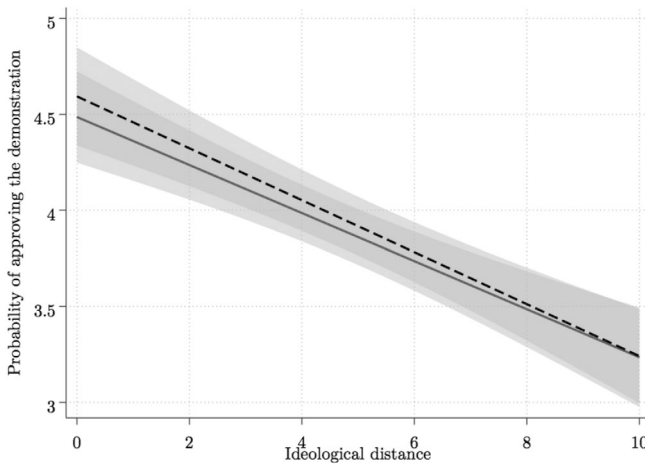


FIGURE 4 Likelihood of authorizing a demonstration and ideological distance in Experiment 1. The solid line captures the control group, while the dashed line refers to the treatment group. Gray ribbons display 95% confidence intervals.

TABLE 2 Regression model: Likelihood of authorizing a demonstration in Experiment 1

Variables	(1) Baseline	(2) Including controls	(3) Including controls	(4) Full set of controls
Absolute distance from organization	-0.125*** (0.021)	-0.123*** (0.021)	-0.118*** (0.020)	-0.126*** (0.020)
Treatment: temporal blind	0.107 (0.177)	0.131 (0.176)	0.093 (0.169)	0.037 (0.170)
Treatment: temporal blind * Abs. distance from organization	-0.010 (0.029)	-0.015 (0.029)	-0.008 (0.028)	0.002 (0.028)
Female		0.029 (0.108)	0.032 (0.104)	0.114 (0.105)
Age		0.009** (0.004)	0.000 (0.004)	0.001 (0.004)
Highly educated		0.053 (0.116)	-0.026 (0.113)	-0.065 (0.113)
Nonurban area		-0.248** (0.111)	-0.176* (0.107)	-0.108 (0.107)
General ideology (liberal- conservative)			-0.057*** (0.020)	-0.056*** (0.020)
Importance of the right to demonstrate			0.137*** (0.023)	0.139*** (0.024)
Active use of the right to demonstrate			0.002 (0.112)	0.010 (0.112)
Intense negative attitude on abortion			-0.025 (0.123)	-0.017 (0.125)
Catholic				-0.193 (0.128)
Constant	4.486*** (0.124)	4.159*** (0.220)	3.679*** (0.282)	3.800*** (0.286)
Observations	481	481	481	466
R ²	0.147	0.168	0.246	0.257

Note: SEs in parentheses. Incremental *F*-test does not provide evidence that the model with an interaction term (column 1) is better than the model without an interaction term ($F = 0.12$, p -value = 0.726). Column 1 presents baseline regression where we only include dummy variables for the experimental groups, in addition to ideological distance. Column 2 extends the model by controlling for basic individual level characteristics. Column 3 further extends the model and controls for attitudinal positions for: “General ideology (liberal-conservative),” “General importance of the right to demonstrate,” “Active use of the right to demonstrate,” “Intense negative attitude on abortion.” For a description of the variables, see Table E1.3 in the Supporting Information.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

does not reach any conventional statistical significance level, suggesting that the pattern of relationship does not vary between the two experimental groups. In other words, we cannot detect heterogeneity of effects across the experimental groups. The examination of the marginal effects confirms this finding.²³ Therefore, we do not find evidence to support the effectiveness of the behavioral intervention in mitigating the ideological bias when deciding whether to approve a demonstration (H_3).

It should be noted that we performed a robustness check to confirm that our experimental manipulation did not influence the respondents' ranking of their ideology. We found no statistically significant differences between the control and treatment groups in their ranking of ideology ($F(1) = 0.88, p = 0.347$), and thus verified that our experimental manipulation did not affect our independent variable.

EXPERIMENT 2: SALIENCY AND ETHICAL DISSONANCE

In this experiment, we have examined H_1 and H_4 . In particular, we first replicated the design of Experiment 1, to examine the existence of ideological bias in the context of constitutional rights (H_1). Next, Experiment 2 was designed to investigate another behavioral intervention with the aim of debiasing the decision makers in the context of constitutional rights, utilizing the insights from the self-awareness theory (H_4).

Experimental design

As before, each participant was asked to imagine that he or she were a public official (mayor) and had to decide whether to approve an application to hold a demonstration. One organization had a conservative ideology and the other a liberal ideology (the same organizations as in Experiment 1). Before reading the application and answering the main question of interest, participants were requested to read, or read and sign, one of two declarations (treatment groups). The first declaration underlined the participants' obligation to be impartial, and the second focused on the importance of giving priority to constitutional rights over other public interests. In order to elicit truthful responses and reduce experimental demand, we informed the participants in the experiment that their decisions would remain anonymous. However, it was also important to make the signature meaningful. Therefore, in the signature groups, we asked the participants, in addition to reading the declaration, to type a sentence summarizing the declaration²⁴ and then to tick a box. This was designed to increase the saliency of their ethical obligation. Again, as in Experiment 1, we assessed each

participant's ideological stance in order to create a measure of "ideological distance" between the participants and the applying organization. Since each experimental group (one control group and four treatment groups) faced both organizations (conservative and liberal), there were 10 experimental groups in total (a 2×5 design). For the experimental groups, see Table E2.1 in the Supporting Information.

Next, participants received the application from one of the two organizations, and the same text as in Experiment 1 explaining the Polish law with respect to the constitutional right to demonstrate and the conditions under which this right can be balanced against other public interests (the exact text can be found in Table E1.1 in the Supporting Information). This information was followed by the question of interest—"What is the probability that you, in your capacity as the mayor of one of the major cities in Poland, would authorize the requested demonstration?" (1–5 Likert scale).

We followed the objective self-awareness theory in our choice of the experimental design. First, we created a stimulus to make the participants focus attention on their "self" and the standard and trigger the objective self-awareness state. The declaration (and the signature) made the standard (impartiality or respecting constitutional rights) very salient, as well as the person's ethical obligation to follow this standard in their imagined position (as a mayor). Next, when the participants had to decide whether to approve an application, the comparison between the self and the standard was automatically prompted. At this stage, they encountered their ideological bias and the resulting desire to favor their ideological group (or disfavor the opposing ideological group). On the other hand, the participants faced the salient standard of being impartial (or giving priority to constitutional rights) and ignoring such personal interests. Participants were expected to experience what can be termed "anticipated ethical dissonance" (Barkan et al., 2015) between their self and the standard. Given the fact that the standard became salient before the decision was made, it was easier for the participants to adjust their behavior to the standard (deciding without ideological bias) rather than trying to change the standard. Therefore, we expected our experimental intervention to reduce the weight that participants accorded to the ideology of the applying organization when determining the probability of approving its application to demonstrate.

Experimental procedure

We used a representative sample (in terms of age, gender, and education) of the Polish population, which was recruited by a professional survey firm, Kantar Polska. The sample included 1952 participants ($N = 1952$), of which about 51% were female, 30% highly educated (bachelor's or master's degree) and 36% non-urban. The average age recorded in the sample was nearly 43 and the average

position on the 0–10 ideological scale regarding abortion was 4.05 (where 0 = pro-choice and 10 = antiabortion). The study was conducted from April 12 to May 9, 2019. The balancing tests across experimental groups and detailed summary statistics of all control variables are provided in Tables E2.2 and E2.3 in the Supporting Information.

Figure 5 shows the distribution of the ideological distance variable. As in Experiment 1, the distribution of this variable has three clear modes. Approximately, 17% of respondents are perfectly aligned (no ideological difference), almost 12% of respondents are in the middle position (a score of 5) and some 16% of respondents are perfectly misaligned (a score of 10). The distribution is quite balanced in the sense that a similar percentage of respondents are located below (45%) and above (44%) the middle position.²⁵

Results

We employ an OLS regression model to test for H_1 and H_4 with the data from Experiment 2. Specifically, we estimate the model as specified in Equation 2.

$$Y_i = \beta_0 + \beta_1 \mathbf{D}_i + \beta_2 \text{Dist}_i + \delta_1 \mathbf{D}_i \times \text{Dist}_i + \varepsilon_i, \quad (2)$$

where Y_i is a dependent variable—that is, the likelihood of approving demonstration. \mathbf{D}_i is a set of dummy variables capturing all four treatments; reference being the control group. Dist_i represents the distance between the applying organization and respondents' position on the abortion ideological question. Lastly, ε_i refers to an error term. In line with H_1 , we expected that coefficient β_2 would be negative. To be more precise, in the control group (no requirement to read or sign a declaration), we expected to find a negative relationship between ideological distance and the likelihood of approving the demonstration. If the behavioral intervention is effective in mitigating the ideological bias (H_4), this would have implications for coefficients β_1 and/or δ_1 .

In Figure 6, we see four comparisons whereby each comparison tests for the difference between the control condition (no additional information was provided) and four experimental conditions with different declaration setups. Note that the control group remains thus unchanged across all the graphs. Table 3 in turn provides formal results stemming from OLS regressions.

Focusing on the control group, it is clear that there is a negative and relatively strong relationship between the likelihood to approve demonstrations and ideological distance. As seen in column 1 of Table 3, the coefficient depicting ideological bias is estimated at -0.12 and is statistically significant at a level of 1%. Thus, for each one-unit increase of ideological distance, the likelihood of authorizing a demonstration declines by 0.12. The intercept of the model in column 1 of Table 3 is estimated at 4.3; this expresses the average likelihood of

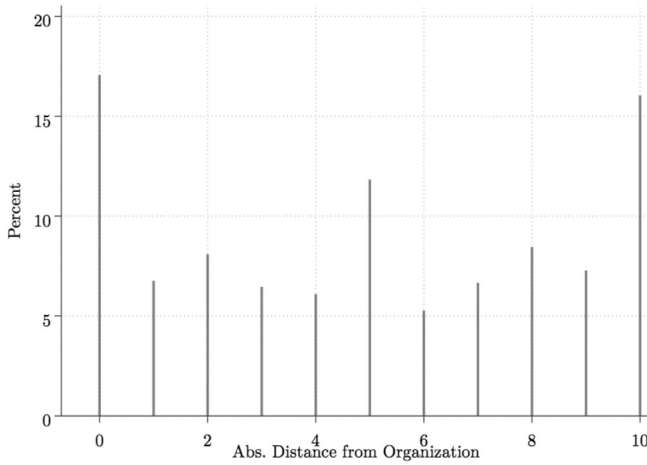


FIGURE 5 Distribution of the ideological distance variable in Experiment 2

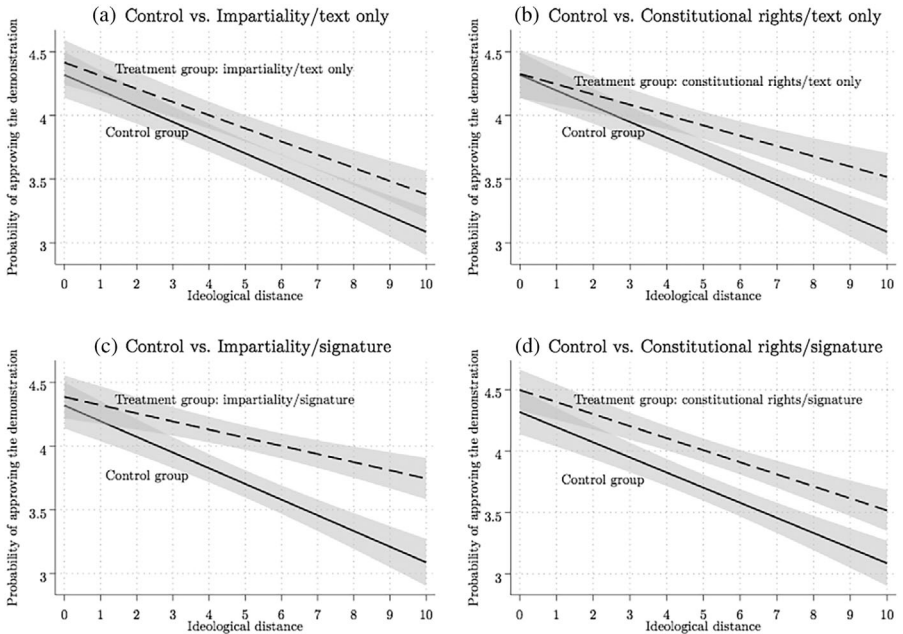


FIGURE 6 Likelihood of authorizing a demonstration and ideological distance in Experiment 2. The solid line captures the control group, while the dashed line refers to various treatment conditions. Gray ribbons display 95% confidence intervals.

approval by respondents who are perfectly ideologically aligned with the applying organization and are assigned to the control condition. The middle respondents (those scoring 5 on the ideological distance scale) approve demonstrations with an average score of approximately 3.7, while the likelihood of demonstration approval in the case of extreme ideological mismatch (10 on the ideological distance scale) is about 3.1. We therefore can conclude that in comparison to middle respondents, ideologically aligned respondents are more likely to authorize a demonstration and ideologically misaligned respondents are less likely to do so. This again is in line with H_1 . The magnitude of the effect (the ideological bias) is similar to what we found in Experiment 1. Given the similarity of the experimental design, these results replicate and indicate the robustness of the ideological bias.

Moving on to comparisons between the control group and the experimental groups, we see that the pattern changes. However, the deviation from the pattern established in the control group varies across experimental groups. Since in none of the models the baseline treatment coefficients (β_1) are statistically significant, indicating null results of treatments for respondents who are fully aligned, we focus our description of results on interpreting the interaction terms (δ_1). When comparing the control group to the experimental group where respondents were subject to *impartiality declaration* treatment (Panel A in Figure 6), we observe a relatively least pronounced deviation from the pattern established in the control group. Even though the interaction term that tests for this deviation is not statistically significant, a good practice in the interaction effect models including treatment dummies and multiscores ordinal variables is to carefully inspect the marginal effects, that is, the effect of treatment (the difference between the treatment and control groups in our case) on all levels of the ordinal variable (ideological distance in our case). In other words, just assessing the coefficient of the interaction term (δ_1 in our case) and its statistical significance is too crude and does not provide sufficient insight into how the effect of one variable is conditioned on the level of another variable (Brambor et al., 2006). Thus, when we look more closely at the differences (marginal effects) between the treatment and control groups at each level of ideological distance (Table 4), we conclude that statistically significant differences do occur at the score of 4 or higher on the ideological distance spectrum (see Figure E2.1 in the Supporting Information).²⁶ Hence, the effect of the impartiality declaration is discernible at the higher levels of ideological distance.

Panel B in Figure 6 displays the differences between the control group and the experimental group where respondents were subject to *constitutional rights declaration* treatment. The disparity between the control group and the treatment group here appears to be wider in comparison to the results of the impartiality declaration treatment. This is reflected in the fact that the interaction term for constitutional rights declaration treatment is statistically significant at the level of 5%. Again, a closer look at the differences between the treatment

TABLE 3 Regression model: Likelihood of authorizing a demonstration in Experiment 2

Variables	(1) Baseline	(2) Including controls	(3) Including controls	(4) Full set of controls
Absolute distance from organization	-0.123*** (0.014)	-0.125*** (0.014)	-0.116*** (0.013)	-0.118*** (0.013)
Impartiality/only text	0.097 (0.125)	0.077 (0.123)	0.043 (0.114)	0.046 (0.116)
Constitutional rights/only text	0.008 (0.126)	-0.002 (0.125)	0.017 (0.116)	0.020 (0.117)
Impartiality/signature	0.067 (0.126)	0.055 (0.124)	0.064 (0.116)	0.062 (0.116)
Constitutional rights/signature	0.180 (0.126)	0.171 (0.124)	0.163 (0.115)	0.163 (0.117)
Impartiality/only text * Abs. distance from organization	0.020 (0.021)	0.019 (0.020)	0.026 (0.019)	0.028 (0.019)
Constitutional rights/only text * Abs. distance from organization	0.042** (0.021)	0.042** (0.020)	0.034* (0.019)	0.034* (0.019)
Impartiality/signature * Abs. distance from organization	0.059*** (0.020)	0.059*** (0.020)	0.053*** (0.019)	0.055*** (0.019)
Constitutional rights/signature * Abs. distance from organization	0.025 (0.021)	0.027 (0.020)	0.024 (0.019)	0.026 (0.019)
Female		-0.183*** (0.047)	-0.128*** (0.044)	-0.159*** (0.045)
Age		0.008*** (0.002)	0.001 (0.002)	0.001 (0.002)
Highly educated		0.203*** (0.050)	0.103** (0.047)	0.103** (0.048)
Nonurban area		-0.045 (0.048)	-0.016 (0.045)	-0.018 (0.046)
General ideology (liberal-conservative scale)			-0.023*** (0.009)	-0.028*** (0.009)

(Continues)

TABLE 3 (Continued)

Variables	(1) Baseline	(2) Including controls	(3) Including controls	(4) Full set of controls
Importance of the right to demonstrate			0.161***	0.158***
			(0.010)	(0.011)
Active use of the right to demonstrate			0.155***	0.146***
			(0.045)	(0.046)
Intense negative attitude on abortion			0.006	-0.020
			(0.050)	(0.051)
Catholic				0.166***
				(0.056)
Constant	4.318***	4.021***	3.072***	3.033***
	(0.088)	(0.118)	(0.141)	(0.145)
Observations	1952	1952	1952	1874
R^2	0.111	0.140	0.255	0.261

Note: SEs in parentheses. Incremental F -test provides evidence that the model with an interaction term (column 1) is better than the model without the interaction terms ($F = 2.42$, p -value = 0.047). Column 1 presents baseline regression where we only include dummy variables for the experimental groups, in addition to ideological distance. Column 2 extends the model by controlling for basic individual level characteristics. Column 3 further extends the model and controls for attitudinal positions for: "General ideology (liberal-conservative)," "General importance of the right to demonstrate," "Active use of the right to demonstrate," "Intense negative attitude on abortion."

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

and the control group at each score of ideological distance indicates that the statistically significant disparity occurs at the score of 4 and higher on the ideological distance spectrum (see Table 5 below and Figure E2.3 in the Supporting Information).²⁷ That is, in both treatments (impartiality and constitutional rights declarations), we observe differences between the treatment and control groups at the higher levels of ideological distances. The effect of the constitutional right declaration seems to be larger in magnitude than that of the impartiality declaration, yet the test for differences between the two does not yield statistically significant results (see the "Comparisons across Treatment Groups" section in the Supporting Information).

The two bottom panels in Figure 6 display the differences between the control group and treatment groups when respondents were required to copy and sign the declaration (and not only read it). The comparison between the *impartiality declaration + signature* group and the control group in Panel C clearly

TABLE 4 Disparity between *impartiality declaration* treatment and control groups (marginal effects) at various levels of ideological distance

Ideological distance	Marginal effect	SEs	p-value	95% confidence intervals	
0	0.097	0.125	0.436	-0.147	0.341
1	0.117	0.109	0.282	-0.096	0.330
2	0.137	0.095	0.149	-0.049	0.323
3	0.156	0.083*	0.060	-0.007	0.320
4	0.176	0.076**	0.020	0.027	0.325
5	0.196	0.074***	0.008	0.051	0.341
6	0.216	0.077***	0.005	0.065	0.367
7	0.236	0.085***	0.006	0.068	0.403
8	0.255	0.098***	0.009	0.064	0.447
9	0.275	0.112***	0.014	0.056	0.495
10	0.295	0.128***	0.021	0.044	0.546

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

TABLE 5 Disparity between *constitutional rights declaration* treatment and control groups (marginal effects) at various levels of ideological distance

Ideological distance	Marginal effect	SEs	p-value	95% confidence intervals	
0	0.008	0.126	0.951	-0.240	0.256
1	0.050	0.110	0.650	-0.166	0.266
2	0.092	0.096	0.335	-0.096	0.280
3	0.135	0.084	0.110	-0.030	0.299
4	0.177	0.076**	0.021	0.027	0.326
5	0.219	0.074***	0.003	0.075	0.364
6	0.261	0.077***	0.001	0.111	0.412
7	0.304	0.085***	0.000	0.137	0.470
8	0.346	0.097***	0.000	0.156	0.536
9	0.388	0.112***	0.001	0.169	0.607
10	0.430	0.128***	0.001	0.179	0.681

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

suggests that ideological bias was mitigated in the treatment group. In this experimental group, the interaction term with the ideological distance is largest in magnitude (compared to the other interaction terms) and statistically significant at a level of 1%. When focusing on particular scores of ideological distance, we already detect statistically significant differences between the treatment

group and control group starting at a score of 3 (see Table 6 and Figure E2.5 in the Supporting Information).²⁸ In comparison to the impartiality declaration treatment that did not require respondents to copy and sign the declaration, the impartiality declaration + signature treatment proved to be statistically more effective, particularly in the case of respondents scoring 5 or more on the ideological distance scale (see the “Comparisons across Treatment Groups” section in the Supporting Information).

Finally, Panel D of Figure 6 displays the ideological distance scores of the *constitutional rights declaration + signature* treatment group in comparison to the scores of the control group. This treatment appears to have the most balanced effect across all ideological distance scores. A closer look at the divergence between the treatment and control groups reveals statistically significant differences only for respondents scoring 2 or higher (see Table 7 below and Figure E2.7 in the Supporting Information).²⁹ Thus, for the most aligned respondents (those scoring 0 or 1 on the ideological distance scale), we cannot provide evidence that the treatment is effective. It is interesting to note that we did not find statistical evidence to support our assumption that adding a requirement to copy and sign the declaration would increase the impact of the constitutional rights declaration treatment (see the “Comparisons across Treatment Groups” section in the Supporting Information).

Based on these results, we conclude that the treatment conditions—debiasing techniques—significantly mitigated the ideological bias in authorizing demonstrations. Thus, we provide evidence for H₄. However, the debiasing effects of treatments were found only for respondents who are ideologically distant from

TABLE 6 Disparity between *impartiality declaration + signature* treatment and control groups (marginal effects) at various levels of ideological distance

Ideological distance	Marginal effect	SEs	<i>p</i> -value	95% confidence intervals	
0	0.067	0.126	0.597	-0.181	0.314
1	0.126	0.110	0.254	-0.090	0.342
2	0.185	0.096*	0.054	-0.003	0.374
3	0.244	0.085***	0.004	0.079	0.410
4	0.303	0.077***	0.000	0.153	0.454
5	0.363	0.074***	0.000	0.218	0.507
6	0.422	0.076***	0.000	0.272	0.571
7	0.481	0.084***	0.000	0.316	0.645
8	0.540	0.095***	0.000	0.353	0.727
9	0.599	0.109***	0.000	0.384	0.814
10	0.658	0.125***	0.000	0.413	0.904

p* < 0.1; *p* < 0.05; ****p* < 0.01.

TABLE 7 Disparity between *constitutional rights declaration + signature* treatment and control groups (marginal effects) at various levels of ideological distance

Ideological distance	Marginal effect	SEs	<i>p</i> -value	95% confidence intervals	
0	0.180	0.126	0.153	-0.067	0.426
1	0.205	0.110*	0.062	-0.010	0.420
2	0.230	0.095**	0.016	0.043	0.417
3	0.255	0.084***	0.002	0.090	0.419
4	0.280	0.076***	0.000	0.130	0.429
5	0.305	0.074***	0.000	0.160	0.449
6	0.330	0.077***	0.000	0.179	0.480
7	0.355	0.085***	0.000	0.188	0.521
8	0.380	0.097***	0.000	0.190	0.570
9	0.405	0.111***	0.000	0.187	0.623
10	0.430	0.127***	0.001	0.180	0.680

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

the applying organization. We did not find evidence that the behavioral interventions are effective in the case of strongly aligned respondents, suggesting that in-group love bias might be particularly ingrained. These results are robust to including control variables (see columns 2 and 3 of Table 3) and dealing only with respondents who passed the attention check (see the “Attention Check” section in the Supporting Information). In the latter case, however, the behavioral intervention works on a narrower group of strongly misaligned respondents.

We also performed a robustness check in Experiment 2 to see whether our experimental manipulation influenced the respondents’ ranking of their ideology. We found no statistically significant differences between the control and treatment groups in their ranking of ideology ($F(4) = 0.76, p = 0.554$), thus confirming that our experimental manipulation did not affect our independent variable.

GENERAL DISCUSSION AND POLICY IMPLICATIONS

In this paper, we have attempted to address two questions. First, we investigated the nature of ideological bias in the context of constitutional rights. Second, we examined different behavioral interventions to assess whether ideological bias can be mitigated. Our findings in both experiments confirmed previous findings on ideological bias. Participants were more prone to approve a demonstration when ideologically aligned with the applying organization. However, contrary

to the specific literature on morality-based groups (Parker & Janoff-Bulman, 2013, pp. 84, 92–93; Weisel & Böhm, 2015, p. 117), we demonstrated that this bias is driven by in-group love, rather than a combination of in-group love and out-group hate. In addition, while previous studies have found that partisanship leads people to assess the risk from opposing groups more harshly (Kahan et al., 2012), results from our control group indicate yet another mechanism at work. In their motivated-reasoned attempt to bring a constitutional decision into accord with their ideology, people can also alter their assessment of the value and importance of basic rights. With respect to the behavioral interventions, we did not find evidence that applying temporary blinding procedure and utilizing cognitive dissonance can mitigate the ideological bias (Experiment 1). One explanation might be that the behavioral intervention was not sufficiently strong to evoke inconsistency (triggering cognitive dissonance) that could only be resolved through a change in behavior. In other words, participants might have separated the two stages in a way that did not induce strong cognitive inconsistency. They could give priority to constitutional rights in stage one and still convince themselves in stage two that despite the importance of these rights, the public interest/risk of public disorder was sufficiently strong in this particular case to render the latter more important. The fact that the demonstrators' identity is to some extent a relevant consideration in determining the probability of public disorder supported this line of thinking. The same argument is valid the other way around: participants could initially rank the importance of constitutional rights as average, but then exhibit a higher likelihood of approving the demonstration after learning that the applying organization is ideologically close to them and thus assuming that it did not pose a threat to public order.

In contrast, the second behavioral intervention, which combined the saliency of morality with cognitive dissonance, was found to be successful to some extent in mitigating ideological bias (Experiment 2). In particular, requiring the participants to read a declaration, or read and sign it, as a mechanism of self-commitment to be impartial in their decision or to prioritize constitutional rights reduced the influence of the ideological bias. When signing such a declaration, participants were more prone to approve the application for demonstration—even by organizations ideologically distant from them—as compared to participants who were not exposed to this experimental manipulation. The relative effectiveness of this intervention, as compared to the behavioral intervention in Experiment 1, might be explained by this intervention's ability to evoke a sense of inconsistency that was harder to mitigate simply through self-justification. By signing the declaration, the respondents became strongly committed to a certain standard of behavior. Allowing themselves to be driven by ideological bias would evoke a strong sense of inconsistency between the actual self and the ideal self and create a negative effect. The easy way to avoid this was to give less weight to their ideological views. In Experiment 1, allowing ideological bias to



shape one's decision on whether to approve a demonstration did not constitute a direct contradiction of the previous behavior (ranking the importance of constitutional rights). In Experiment 2, it did.

Despite the promising results in Experiment 2, it should be noted that ideological bias was mitigated, but not entirely eliminated. Our results did not support the effectiveness of the intervention in reducing or mitigating the ideological bias in cases of complete ideological alignment between the participants and the applying organization. This might support our findings in Experiment 1, where the ideological bias was in particularly strong among the most aligned participants. These results suggest that in-group love is an extremely powerful bias that is difficult to eliminate. Nevertheless, it should be noted that based on our sample size, we were able to detect medium to large effect sizes of interventions. Therefore, we cannot exclude the possibility that a small sized effect of the intervention exists.

The paper offers several contributions to existing literature. First, it reveals the robust influence of ideology on decisions about basic rights through the effect it has on the value people attributed to the right as well as the assessment of danger posed by opposing ideological group. Second, to the best of our knowledge, we are the first to identify the nature of ideological bias in the context of constitutional rights. As explained in the introduction, the magnitude of this bias varies depending on whether it is driven by in-group love or out-group hate. Moreover, the search for potential solutions should be informed by this question. To a certain extent, our findings seem to challenge recent literature asserting that despite the general dominance of in-group love, there is a preponderance of out-group hate in the case of morality-based groups (Parker & Janoff-Bulman, 2013). We chose typical morality-based groups (anti-abortion versus pro-choice ideology). In this case, group identity is interdependently influenced by the in-group and the out-group, because the out-group threatens the in-group's perception of right and wrong. Therefore, our results are conservative and contest the contention that out-group derogation is an equally powerful motivation for the intergroup bias in such cases (Weisel & Böhm, 2015, p. 117).

Third, despite the extensive literature on ideological bias, empirical testing of potential solutions is scarce. Therefore, the additional novelty of this paper is to examine different behavioral interventions to address this problem.

The use of a representative sample and a randomized survey experiment increased the strength of the results. Furthermore, most of the studies on ideological bias have been conducted on Anglo-American participants. Therefore, an additional contribution of this paper is to expand the target population for investigating ideological bias in constitutional questions and intergroup bias. Finally, contrary to the pessimistic view that e-signatures do not possess the power of real signatures to curb dishonest behavior (Chou, 2015), we demonstrated that e-signatures are also powerful in enhancing ethical behavior.

In terms of policy implications, our study offers a relatively cheap intervention in procedures where constitutional rights need to be balanced against other interests. In light of the different findings in Experiments 1 and 2, we also suggest ensuring that the intervention is sufficiently strong to evoke dissonance (when cognitive dissonance is used as part of the behavioral intervention) and is on the same dimension where the clear strategy for reduction is the avoidance of the ideological bias influence. In other words, when using cognitive dissonance as the psychological mechanism to induce certain behavior, it is important that the design of the procedural rule really evokes such dissonance. Furthermore, the solution to resolve this dissonance should be the adoption of the desired behavior. Therefore, for example, policy makers could consider introducing a mandatory procedure where public officials sign declarations before making their decisions and commit themselves to be impartial. Such procedure is not costly to implement and may have an important impact even if the effect is limited.

Despite the contribution of our experiments, they also have several limitations, which open an opportunity for further research. First, our studies are embedded in the theoretical framework and literature on the intergroup bias (in-group love and out-group hate). However, an alternative explanation for the ideological bias which is found in the context of approving demonstrations might be promotion of own ideological commitments. In other words, our participants might have decided to approve a demonstration of an ideologically aligned organization or reject a demonstration by ideologically misaligned organization for the reason of promoting “legalization” or prohibition of abortions (depending on their stance). This is in contrast to being in favor or against a particular group, as is sometimes the case with racial or ethnical groups for example. Our experimental designs do not allow to neatly disentangle those two motivations. Nevertheless, in Experiment 1, we found that participants approved to a larger extent a demonstration by an ideologically aligned organization, but at the same time, did not treat differently the ideologically misaligned organization (what we have termed as finding in-group love but not out-group hate). This might suggest that the sole promotion of own ideological interests is not the strongest explanation here. Since both, approving a demonstration by an ideologically aligned organization and rejecting a demonstration by an ideologically misaligned organization, promotes the interests of the participants, we would expect to find both types of decisions (favoritism and derogation). In any case, we would call for future research on ideological bias to try to specifically disentangle those two decision channels (inter-group bias versus promotion of own ideological commitments).

Second, the study was conducted on the Polish general population. Though we believe this context was suitable in light of the current political polarization, additional research should be conducted in other countries in order to generalize our results. Furthermore, the participants were lay people rather than real



policy/decision-makers; therefore, the next step is to test our questions on experts. Nevertheless, many studies have shown that experts and lay people act similarly when behavioral biases are involved (Guthrie et al., 2000, Plous, 1993, p. 258). In addition, even though our participants were not real policy makers, we should remember that an experimental setting allows for more freedom than a real setting. The decisions of policy makers may still be challenged in court. The respondents in our study participated in a completely anonymous setting and had no stakes in fulfilling or renegeing on their commitment. Therefore, the effect of the interventions in Experiment 2 might even be considered conservative.

Finally, as with other studies on behavioral interventions, our paper presents a one-shot game. It is unclear whether the mitigating effect on ideological bias can be maintained in the long run with the same intervention. The long-term effect of such interventions should be studied. Nevertheless, in view of the low costs of the behavioral intervention, policy makers should consider implementing it.

ACKNOWLEDGMENTS

Helpful suggestions by Mota Kremnitzer, Raanan Sulitzeanu-Kenan, Yuval Feldman, Ilana Ritov, Liat Netzer, Marina Motsenok, Ori Plonsky, Netta Barak-Corren, audiences at 15th Conference on Empirical Legal Studies (March 2022), participants at the workshop on “Should Data Shape Private Law: Between Stereotypes and Personalization” (June 2020), and anonyms referees are gratefully acknowledged.

FUNDING INFORMATION

The research for this article was undertaken as part of the “Proportionality in Public Policy” project at the Israel Democracy Institute, with support from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007–2013), ERC grant no. 324182.

DATA AVAILABILITY STATEMENT

Data necessary to replicate the results of this article are available upon request from the corresponding author.

ENDNOTES

- ¹ The principle was first emphasized in the Universal Declaration on Human Rights in 1948 in the Assembly, UN General. 1948. Universal Declaration of Human Rights. *UN General Assembly*, 302(2) and has been reiterated in many national constitutions, international conventions and judicial decisions.
- ² Since 2015, when the Law and Justice (PiS) party won the parliamentary elections and started governing, among others, the judicial independence in general and in the Polish Constitutional Tribunal (TK) in particular was seriously challenged. The actions undertaken by PiS aimed to

pack the TK with its own appointees and paralyze its functioning, at least until PiS would secure a safe majority in the court. For a detailed description of these challenges, see Sadurski (2019) and Garlicki (2019). On the general politicization of the Polish TK, see Kantorowicz and Garoupa (2016). Overall, it is well documented that Poland has observed a gradual drop in constitutional compliance over the last five years. See a novel dataset on constitutional compliance by Gutmann et al. (2022).

- ³ For India, see Chandrachud (2013); for Israel, see Barak (2012); for Hong Kong, see Yin (2014), pp.281-312.
- ⁴ For Canada, see Berger (2010); for Brazil, see Neto (2018).
- ⁵ For South Africa, see *S v. Makwanyane & Another* 1995 (3) SA 391 (CC) at para. 104 (S. Afr.).
- ⁶ See a detailed empirical analysis of the implementation of the proportionality principle by courts in Germany, Canada, South Africa, Israel, Poland, and India in Kremnitzer et al. (2020) and Petersen (2017).
- ⁷ Some scholars do not include the first stage within the application of the principle, see Alexy (2010, pp. 397–414).
- ⁸ The restriction on a public gathering scheduled for the same time and place as another, cyclical gathering, has been challenged by the Polish Ombudsman (“Commissioner for Human Rights”) and others as unconstitutional. See the report by the Ombudsman, *Wolność zgromadzeń w Polsce w latach 2016–2018*, Kolakowska-Skorupa & Wojsyk (2018).
- ⁹ In the appeal, the organizers argued that the grounds for limiting the freedom of assembly should be interpreted very narrowly. In case of conflict of constitutional principles, freedom of assembly is equivalent to security and public order. They further argued that police provided general updated statistics, which were irrelevant to the case at hand. They also noted that municipal authorities did not consider the possibility of eliminating potential threats by increasing security measures. The organizers also contended that the municipality’s ban on the gathering was driven by extralegal circumstances, namely pressures from rival groups.
- ¹⁰ See examples of research that found strong effects of racial or ethnic in-group bias in Rachlinski et al. (2008), Shayo and Zussman (2011), and Gazal-Ayal and Sulitzeanu-Kenan (2010).
- ¹¹ One might suggest that a full blinding procedure should have been maintained, for example, by not exposing the participants to the identity of the organization, and just giving them information on what the police thinks the risk is to public order. However, in the center of the study is the individual evaluation of the potential risk, in light of one’s ideological bias. Giving direct information to the participants about what exactly the police might consider to be the risk, would miss the exact aspect we are investigating—whether interpretation of a *vague* risk is influenced by the potential ideological bias of the decision maker. Moreover, we were interested in examining a solution which can be implemented in reality in order to have policy recommendation from our empirical work. Blinding entirely the approval procedure of demonstrations by requiring the decision maker to only rely on the estimated risk by the police, in fact delegates the decision (together with the potential problem of ideological bias) to the police. Therefore, we think it was important to keep the design as temporary blinding procedure.
- ¹² A dictator game design refers to an experimental setting in which two participants are paired and have a “common pool” of money (either endowment or earned). One of the participants then decides how to allocate the money in this pool. The second participant can only accept or reject the allocation. If the offer of the “dictator” is rejected, the participants do not receive anything from the “common pool.”
- ¹³ A well-known study by Shu et al. (2012), which was later retracted by the journal Proceedings of the National Academy of Sciences (PNAS) due to evidence of manipulation of data, applied insights from objective self-awareness theory to address the problem of dishonest reporting, using



the manipulation of the location of a person's signature, which they found to be effective. It should be stressed that despite the relevance of that study, we derived our predictions from the theory of objective self-awareness, rather than the results of that study. Furthermore, we did not compare signing at the beginning to signature at the end, but rather different forms of declarations, with or without a signature as compared to no prior form of commitment.

- ¹⁴ It is important to note that our experiments were conducted prior to a controversial judgment by the Polish TK (K 1/20), which limited the legality of abortion.
- ¹⁵ The antiabortion organization was the Polish Association for the Protection of Human Life (Polskie Stowarzyszenie Obrońców Życia Człowieka) and the pro-choice organization was the Federation for Woman and Family Planning (Federacja na rzecz Kobiet i Planowania Rodziny).
- ¹⁶ Participants were asked to answer the following question: *How would you define your views on abortion, on a scale of 0–10, where 0 = (liberal) I am in favor of free choice on abortion and 10 = (conservative) I am strongly against abortion?* For each of the experiments we have measured whether the treatment had an effect on the answer to this question given that it was asked after the main questions. In none of the studies such an effect was found. Thus, confirming the reliability of our measurement of ideological distance. More details on the tests are in the results sections.
- ¹⁷ Participants were asked to answer the following question Q1: *To what extent, in your opinion, should a possible risk to public order and safety preclude demonstrations?* (1–5 Likert scale).
- ¹⁸ Participants were asked to answer the following question Q2: *What is the probability that you, in your capacity as the mayor of one of the major cities in Poland, would authorize the requested demonstration?* (1–5 Likert scale).
- ¹⁹ Another way to explain the expected decision mechanism could be through the constraint satisfaction framework (Holyoak & Simon, 1999; Simon et al., 2004). Because both tasks involve similar underlying issues (rating importance of constitutional rights, and then deciding on the provision of a constitutional right), the activation of the first assessment (importance) may spread coherence to the second assessment (approval probability). This in our context would mean that rating higher/lower the importance of the constitutional right to demonstrate, should also lead to higher/lower probability of approving the demonstration, irrespective of the ideological distance.
- ²⁰ There is no evidence that the ideological distance variable does not balance across the two treatment and control experimental groups ($t = -1.41$, p -value = 0.159).
- ²¹ Our main results are based on the sample that contains subjects who failed the attention check. (For more details about the attention check in the context of this study, see the Attention Check section in the Supporting Information). We follow here Aronow et al. (2019) who show that the practice of discarding subjects who failed manipulation/attention checks leads to significant bias and argue for a focus on what is revealed without dropping respondents. Nonetheless, the Supporting Information—which presents the results for a trimmed sample (that is, when the respondents who failed the attention check are excluded)—indicates that the results do not change significantly.
- ²² Since the regression coefficients do not enable direct verification of the differences between the treatment and control groups at the misalignment extreme, we provide in Table E1.4 and Figure E1.1 (Supporting Information) the differences in weighting the importance of the right to demonstrate between the treatment and control groups for each level of ideological misalignment, which is a standard procedure for the interaction effect models. Finally, in line with recently devised procedure, we also test if estimating interaction effects with linear model is plausible. The Wald test informs that we cannot reject the null hypothesis that linear model is plausible (p -value = 0.606), see Hainmueller et al. (2019). It is to note that we have also performed an equivalence test to show that with our sample size we were able to detect a medium effect size at

the maximum ideological distance of 10 where the difference was not statistically significant (Figure E1.2 in the Supporting Information). The equivalence test employed here is based on Lakens et al. (2018).

- ²³ Table E1.5 and Figure E1.3 (Supporting Information) confirms this as it demonstrates that the differences between the treatment and control groups are not statistically significant at any level of ideological distance. Furthermore, according to an equivalence test for ideological distance of 0, 5, and 10, we show that with our sample size we were able to detect a medium effect size (Figure E1.4 in the Supporting Information). Also here the Wald test informs that we cannot reject the null hypothesis that linear model is plausible (p -value = 0.515).
- ²⁴ In the setting of the programmed experiment, participants were prevented from simply copying and pasting the sentence. They had to type the sentence.
- ²⁵ There is no evidence that the ideological distance variable does not balance across the experimental groups ($F = 0.33$, p -value = 0.857).
- ²⁶ According to an equivalence test performed for respondents with the ideological distance of 0 (fully aligned respondents), we show that with our sample size we were able to detect a medium effect size of treatment (Figure E2.2 in the Supporting Information). The Wald test in turn informs us that we cannot reject the null hypothesis that linear model is plausible (p -value = 0.221).
- ²⁷ By employing an equivalence test, we show that with our sample size we were able to detect a medium effect size of treatment for the fully aligned respondents, that is, respondents with the ideological distance of 0 (Figure E2.4 in the Supporting Information). Furthermore, in line with the Wald test, we cannot reject the null hypothesis that linear model is plausible (p -value = 0.731).
- ²⁸ The equivalence test demonstrates that with our sample size we were able to detect a large effect size and the upper bounds of the medium effect sizes of treatment at the ideological distance of 0 (Figure E2.6 in the Supporting Information). The Wald test informs that we cannot reject the null hypothesis that linear model is plausible (p -value = 0.080).
- ²⁹ The equivalence test indicates that with our sample size we were able to detect large effect sizes of treatment for the fully aligned respondents, that is, those with ideological distance of 0 (Figure E2.8 in the Supporting Information). Moreover, the Wald test informs that we cannot reject the null hypothesis that linear model is plausible (p -value = 0.499).

REFERENCES

- Alexy, R. (2010). *A theory of constitutional rights*. Oxford University Press.
- Allport, G. W., Clark, K., & Pettigrew, T. (1954). *The nature of prejudice*. Addison-Wesley.
- Andenas, M., & Zleptnig, S. (2007). Proportionality and balancing in WTO law: a comparative perspective. *Cambridge Review of International Affairs*, 20, 71–92.
- Aronow, P. M., Baron, J., & Pinson, L. (2019). A note on dropping experimental subjects who fail a manipulation check. *Political Analysis*, 27(4), 572–589.
- Babcock, L., Loewenstein, G., & Issacharoff, S. (1997). Creating convergence: Debiasing biased litigants. *Law & Social Inquiry*, 22(4), 913–925.
- Balliet, D., Wu, J., & De Dreu, C. K. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin*, 140(6), 1556–1581.
- Barak, A. (2012). *Proportionality: Constitutional rights and their limitations*. Cambridge University Press.
- Barkan, R., Ayal, S., & Ariely, D. (2015). Ethical dissonance, justifications, and moral behavior. *Current Opinion in Psychology*, 6, 157–161.

- Berger, B. L. (2010). Section 1, constitutional reasoning, and cultural difference: Assessing the impacts of *Alberta v. Hutterian Brethren of Wilson Colony*. *Supreme Court Law Review*, 51 (2D), 25–46.
- Bolsen, T., Druckman, J. N., & Cook, F. L. (2014). The influence of partisan motivated reasoning on public opinion. *Political Behavior*, 36(2), 235–262.
- Brambor, T., Clark, W. R., & Golder, M. (2006). Understanding interaction models: Improving empirical analyses. *Political Analysis*, 14(1), 63–82.
- Brehm, J. W., & Cohen, A. R. (1962). *Explorations in cognitive dissonance*. John Wiley & Sons Inc.
- Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate? *Journal of Social Issues*, 55(3), 429–444.
- Carter, A. (2016). Proportionality in Australian constitutional law: Towards transnationalism? *Heidelberg Journal of International Law*, 76, 951–966.
- Chandrachud, A. (2013). *Wednesbury reformulated: Proportionality and the supreme court of India*. *Oxford University Commonwealth Law Journal*, 13(1), 191–208.
- Chou, E. Y. (2015). What's in a name? The toll e-signatures take on individual honesty. *Journal of Experimental Social Psychology*, 61, 84–95.
- Choudhry, S. (2006). So what is the real legacy of Oakes? Two decades of proportionality analysis under the Canadian Charter's Section 1. *Supreme Court Law Review*, 35, 501–535.
- Cohen, G. L. (2003). Party over policy: The dominating impact of group influence on political beliefs. *Journal of Personality and Social Psychology*, 85(5), 808–822.
- Cohen-Eliya, M., & Porat, I. (2011). Proportionality and the culture of justification. *The American Journal of Comparative Law*, 59(2), 463–490.
- Cooper, J., & Hogg, M. A. (2007). Feeling the anguish of others: A theory of vicarious dissonance. *Advances in Experimental Social Psychology*, 39, 359–403.
- Cope, K. L., & Crabtree, C. (2020). A nationalist backlash to international refugee law: Evidence from a survey experiment in Turkey. *Journal of Empirical Legal Studies*, 17(4), 752–788.
- Donnelly, J. (2007). The relative universality of human rights. *Human Rights Quarterly*, 29, 281–306.
- Duval, S., & Wicklund, R. A. (1972). *A theory of objective self-awareness*. New York Academic Press.
- Epstein, L., Parker, C. M., & Segal, J. A. (2018). Do justices defend the speech they hate? An analysis of in-group bias on the US supreme court. *Journal of Law and Courts*, 6(2), 237–262.
- Festinger, L. (1962). *A theory of cognitive dissonance*. Stanford University Press.
- Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *The Journal of Abnormal and Social Psychology*, 58, 203–210.
- Furgeson, J. R., Babcock, L., & Shane, P. M. (2008). Do a law's policy implications affect beliefs about its constitutionality? An experimental test. *Law and Human Behavior*, 32(3), 219–227.
- Garlicki, L. (2019). Constitutional court and politics. The polish crisis. In C. Landfried (Ed.), *Judicial power. How constitutional courts affect political transformations* (pp. 141–162). Cambridge University Press.
- Gazal-Ayal, O., & Sulitzeanu-Kenan, R. (2010). Let my people go: Ethnic in-group bias in judicial decisions—Evidence from a randomized natural experiment. *Journal of Empirical Legal Studies*, 7(3), 403–428.
- Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of “blind” auditions on female musicians. *American Economic Review*, 90(4), 715–741.
- Goren, P., Federico, C. M., & Kittilson, M. C. (2009). Source cues, partisan identities, and political value expression. *American Journal of Political Science*, 53(4), 805–820.
- Guthrie, C., Rachlinski, J. J., & Wistrich, A. J. (2000). Inside the judicial mind. *Cornell Law Review*, 86, 777–811.
- Gutmann, J., Metelska-Szaniawska, K., & Voigt, S. (2022). The comparative constitutional compliance database. Working paper, available at. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4035701

- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814–834.
- Hainmueller, J., Mummolo, J., & Xu, Y. (2019). How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice. *Political Analysis*, 27(2), 163–192.
- Halevy, N., Bornstein, G., & Sagiv, L. (2008). “In-group love” and “out-group hate” as motives for individual participation in intergroup conflict: A new game paradigm. *Psychological Science*, 19(4), 405–411.
- Halevy, N., Weisel, O., & Bornstein, G. (2012). “In-group love” and “out-group hate” in repeated interaction between groups. *Journal of Behavioral Decision Making*, 25(2), 188–195.
- Harbo, T. I. (2010). The function of the proportionality principle in EU law. *European Law Journal*, 16(2), 158–185.
- Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology*, 53, 575–604.
- Holyoak, K. J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General*, 128(1), 3–31.
- Huber, G. A., & Malhotra, N. (2017). Political homophily in social relationships: Evidence from online dating behavior. *The Journal of Politics*, 79(1), 269–283.
- Huscroft, G., Miller, B. W., & Webber, G. (Eds.). (2014). *Proportionality and the rule of law: Rights, justification, reasoning*. Cambridge University Press.
- Jackson, V. C. (2014). Constitutional law in an age of proportionality. *Yale Law Journal*, 124, 3094.
- Jolls, C., & Sunstein, C. R. (2006). Debiasing through law. *The Journal of Legal Studies*, 35(1), 199–242.
- Kahan, D. (2010). Fixing the communications failure. *Nature*, 463(7279), 296–297.
- Kahan, D. M., Hoffman, D. A., Braman, D., & Evans, D. (2012). They saw a protest: Cognitive illiberalism and the speech-conduct distinction. *Stanford Law Review*, 64, 851.
- Kantorowicz, J., & Garoupa, N. (2016). An empirical analysis of constitutional review voting in the Polish constitutional tribunal, 2003–2014. *Constitutional Political Economy*, 27(1), 66–92.
- Kořalkowska-Skorupa, M., & Wojsyk, J. (2018). *Wolność zgromadzeń w Polsce w latach 2016–2018*. Biuro Rzecznika Praw Obywatelskich.
- Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review*, 90(4), 1072–1091.
- Kraft, P. W., Lodge, M., & Taber, C. S. (2015). Why people “don’t trust the evidence” motivated reasoning and scientific beliefs. *The Annals of the American Academy of Political and Social Science*, 658(1), 121–133.
- Kremnitzer, M., Steiner, T., & Lang, A. (Eds.). (2020). *Proportionality in action: Comparative and empirical perspectives on the judicial practice* (Vol. 22). Cambridge University Press.
- Kumm, M. (2007). Political liberalism and the structure of rights: On the place and limits of the proportionality requirement. In G. Pavlakos (Ed.), *Law, rights and discourse: The legal philosophy of Robert Alexy* (Vol. 11). Hart Publishing.
- Kumm, M., & Walen, A. D. (2014). In G. Huscroft, B. W. Miller, & G. Webber (Eds.), *Human dignity and proportionality: Deontic pluralism in proportionality and the rule of law: Rights, justification, reasoning* (pp. 67–89). Cambridge University Press.
- Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, 1(2), 259–269.
- Leeper, T. J., & Slothuus, R. (2014). Political parties, motivated reasoning, and public opinion formation. *Political Psychology*, 35, 129–156.
- Mummendey, A., & Otten, S. (1998). Positive–negative asymmetry in social discrimination. *European Review of Social Psychology*, 9(1), 107–143.
- Neto, J. A. (2018). *Borrowing Justification for proportionality*. Springer.

- Parker, M. T., & Janoff-Bulman, R. (2013). Lessons from morality-based social identity: The power of outgroup “hate,” not just ingroup “love”. *Social Justice Research, 26*(1), 81–96.
- Petersen, N. (2017). *Proportionality and judicial activism: Fundamental rights adjudication in Canada*. Cambridge University Press.
- Plous, S. (1993). *The psychology of judgment and decision making*. McGraw-Hill Book Company.
- Pulido, C. B. (2013). The migration of proportionality across Europe. *New Zealand Journal of Public and International Law, 11*(3), 483–515.
- Rachlinski, J. J., Johnson, S. L., Wistrich, A. J., & Guthrie, C. (2008). Does unconscious racial bias affect trial judges. *Notre Dame Law Review, 84*, 1195–1246.
- Rachlinski, J. J., & Jourden, F. (1998). Remedies and the psychology of ownership. *Vanderbilt Law Review, 51*, 1541.
- Rand, D. G., Pfeiffer, T., Dreber, A., Sheketoff, R. W., Wernerfelt, N. C., & Benkler, Y. (2009). Dynamic remodeling of in-group bias during the 2008 presidential election. *Proceedings of the National Academy of Sciences, 106*(15), 6187–6191.
- Rogowski, J. C., & Sutherland, J. L. (2016). How ideology fuels affective polarization. *Political Behavior, 38*(2), 485–508.
- Sadurski, W. (2019). *Poland’s constitutional breakdown*. Oxford University Press.
- Sah, S., Robertson, C. T., & Baughman, S. B. (2015). Blinding prosecutors to defendants’ race: A policy proposal to reduce unconscious bias in the criminal justice system. *Behavioral Science & Policy, 1*(2), 69–76.
- Schulz, K. F., & Grimes, D. A. (2002). Blinding in randomised trials: Hiding who got what. *The Lancet, 359*(9307), 696–700.
- Shayo, M., & Zussman, A. (2011). Judicial ingroup bias in the shadow of terrorism. *The Quarterly Journal of Economics, 126*(3), 1447–1484.
- Shu, L. L., Mazar, N., Gino, F., Ariely, D., & Bazerman, M. H. (2012). Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. *Proceedings of the National Academy of Sciences, 109*(38), 15197–15200.
- Simon, D., Krawczyk, D. C., & Holyoak, K. J. (2004). Construction of preferences by constraint satisfaction. *Psychological Science, 15*(5), 331–336.
- Śledzińska-Simon, A. (2019). *Analiza proporcjonalności ograniczeń konstytucyjnych praw i wolności: teoria i praktyka*. E-Wydawnictwo, Prawnicza i Ekonomiczna Biblioteka Cyfrowa. Wydział Prawa, Administracji i Ekonomii Uniwersytetu Wrocławskiego.
- Śledzińska-Simon, A. (2020). Proportionality analysis by the Polish Constitutional Tribunal. In M. Kremnitzer, T. Steiner, & A. Lang (Eds.), *Proportionality in action: Comparative and empirical perspectives on the judicial practice* (pp. 385–457). Cambridge University Press.
- Sood, A. M., & Darley, J. M. (2012). The plasticity of harm in the service of criminalization goals. *California Law Review, 100*(5), 1313–1358.
- Sulitzeanu-Kenan, R., Kremnitzer, M., & Alon, S. (2016). Facts, preferences, and doctrine: An empirical analysis of proportionality judgment. *Law & Society Review, 50*(2), 348–382.
- Sulitzeanu-Kenan, R., Sheffer, L., Steiner, T. & Barzilai, S. (2022). Does civil rights discourse moderate or escalate political bias? Experimental evidence from Israel and Canada (Unpublished, copy on file with authors).
- Sweet, A. S., & Mathews, J. (2008). Proportionality balancing and global constitutionalism. *Columbia Journal of Transnational Law, 47*, 72.
- Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science, 50*(3), 755–769.
- Tajfel, H. (1981). *Human groups and social categories*. Cambridge University Press.
- Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology, 1*(2), 149–178.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

- Weisel, O., & Böhm, R. (2015). "Ingroup love" and "outgroup hate" in intergroup conflict between natural groups. *Journal of Experimental Social Psychology*, 60, 110–120.
- Yamagishi, T., & Mifune, N. (2009). Social exchange and solidarity: In-group love or out-group hate? *Evolution and Human Behavior*, 30(4), 229–237.
- Yin, L. P. (2014). *The judicial construction of Hong Kong's basic law: Courts, politics, and society after 1997*. Hong Kong University Press.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kantorowicz-Reznichenko, E., Kantorowicz, J., & Weinshall, K. (2022). Ideological bias in constitutional judgments: Experimental analysis and potential solutions. *Journal of Empirical Legal Studies*, 1–42. <https://doi.org/10.1111/jels.12323>