PNAS

Coordination and expertise foster legal textualism

Ivar R. Hannikainen[a,1] (ID), Kevin P. Tobia[b] (ID), Guilherme da F. C. F. de Almeida[c] (ID), Noel Struchiner[d], Markus Kneer[e] (ID), Piotr Bystranowski[f] (ID),
Vilius Dranseika[f] (ID), Niek Strohmaier[g] (ID), Samantha Bensinger[c], Kristina Dolinina[h], Bartosz Janik[i] (ID), Eglė Lauraitytė[h], Michael Laakasuo[j] (ID),
Alice Liefgreen[k] (ID), Ivars Neiders[l] (ID), Maciej Próchnicki[f], Alejandro Rosas[m] (ID), Jukka Sundvall[j], and Tomasz Żuradzki[f] (ID)

A cross-cultural survey experiment revealed a dominant tendency to rely on a rule's letter over its spirit when deciding which behaviors violate the rule. This tendency varied markedly across ($k = 15$) countries, owing to variation in the impact of moral appraisals on judgments of rule violation. Compared with laypeople, legal experts were more inclined to disregard their moral evaluations of the acts altogether and consequently exhibited stronger textualist tendencies. Finally, we evaluated a plausible mechanism for the emergence of textualism: in a two-player coordination game, incentives to coordinate in the absence of communication reinforced participants' adherence to rules' literal meaning. Together, these studies (total $n = 5,794$) help clarify the origins and allure of textualism, especially in the law. Within heterogeneous communities in which members diverge in their moral appraisals involving a rule's purpose, the rule's literal meaning provides a clear focal point—an identifiable point of agreement enabling coordinated interpretation among citizens, lawmakers, and judges.

moral judgment | legal decision making | coordination | cross-cultural research

All 50 US states have passed zero-tolerance alcohol consumption laws, which severely sanction any person below age 21 who drives with detectable alcohol in their bloodstream. In most cases, when these circumstances obtain, the purpose that gave rise to the law—of protecting other road users and saving lives—has also been jeopardized (1). Yet legal rules fall short of perfect sensitivity and specificity. For instance, a driver under the influence of a chemically distinct narcotic, such as ecstasy, could pose a larger threat to road safety. Call this an underinclusion case; the law's literal formulation fails to proscribe an act that undermines the law's spirit. Similarly, some innocuous behaviors, such as rinsing with an alcohol-based mouthwash, might result in a positive test result without elevating the risk of an accident. Call this an overinclusion case; the law's letter proscribes an act that in fact complies with its spirit.

When evaluating these acts on moral grounds, it is abundantly clear whose behavior is worse: we condemn the first agent's reckless conduct and exonerate the second. This capacity arises early in development (2), as children abandon the uncritical submission to authority and autonomously reason about deeper ethical principles (3, 4), and plausibly implicates outcome-based reasoning over the probability and magnitude of harm (5, 6).

Now consider a different question: which of these behaviors violates zero-tolerance laws? Is it the first, which jeopardizes the law's deeper purpose of saving lives (7, 8), or the second, which conflicts with its literal meaning (9, 10)? By pitting the spirit of the law against its letter, these atypical and controversial cases have historically inspired sustained litigation (11) and provide a rare window into the cognitive processes that underlie legal reasoning.

Recent research has established that laypeople view overinclusion cases (proscribed by the letter of the law) as unlawful, despite their innocuity and compliance with the law's spirit (12–14). In turn, they view underinclusion cases (that jeopardize the law's aims) as lawful as long as they comply with its letter. A tendency toward textualist interpretation arises equally in reaction to everyday transgressions of nonlegal rules, such as a rule that prohibits shoes in the house to foster cleanliness (14). A guest in muddy socks is considered to abide by the household rule—whereas a guest who tries on pristine dress shoes is not. This pattern accords with a prevailing stance among legal theorists (15): as a leading textualist scholar puts it, "texts should be taken at face value—with no implied extensions of specific texts or exceptions to general ones—even if the legislation will then have an awkward relationship to the apparent background intention or purpose that produced it." ((16), p. 428). This emphasis on text prevails also in the US court system, where textualism has grown to be a dominant theory of legal interpretation (17). What could lead jurors to disregard their moral reasoning and prioritize the literal scope of a rule when assessing an act's legality?

Significance

The transition from deference to authority to autonomous reasoning is a major landmark in moral development. In this light, it is interesting how citizens and especially legal experts often heed the letter of the law in detriment of their moral standards during judicial decision making. Despite substantial cultural variability in this phenomenon, our study documented a global tendency toward such "textualist" interpretation and provided an explanation for why it might prevail: prioritizing the letter of the law over its spirit helps citizens and judges reach a shared understanding of law's scope, which plausibly brings about long-term social benefits and outweighs the occasional moral cost of adopting a textualist strategy.

One possibility is that adherence to the rule's letter serves as a heuristic when—as in most naturalistic contexts—the rule's spirit is undisclosed, unclear, or unsettled. Even a simple rule, like "No food in the classroom," might admit of many purposes: maintaining cleanliness, minimizing distraction, and/or avoiding student allergies. Therefore, judging a target act by asking whether it undermines the rule's presumed purpose(s) can be impractical and rife with uncertainty. This perspective raises the possibility that the rule's text plays a heuristic role (18, 19), i.e., to offer a cognitively frugal means by which—with a minimal cost to accuracy (i.e., specificity and sensitivity)—individuals may determine which behaviors violate the rule's purpose. Previous evidence, however, casts doubt on this explanation: participants' textualist tendencies persisted even when revealing the rule's purpose and rendering the act's outcomes easily evaluable (14).

In the present work, we pursue a distinct explanation for the emergence of textualism. We conceptualize statutory interpretation as a social dilemma in which individual judges can have mixed motives (20, 21). In a standard mixed-motive game, multiple drivers are approaching an intersection. Each driver has both (i) an individual preference to drive rather than yield to other drivers and (ii) a stronger interest in coordinating with other drivers to avoid a collision. This coordination goal will lead drivers to converge on a Nash equilibrium.

This incentive structure can help us model the context of statutory interpretation: when applying rules to ambiguous or controversial cases, judges may favor conflicting resolutions of the same case—due, in part, to their divergent moral preferences (22). For instance, in the simplest case involving only two judges (see Table 1), judge 1 has a preference to acquit the defendant (and receives a payoff of P_1 from satisfying this private preference) while judge 2 has a preference to convict them (receiving the payoff P_2 in that case). This model can be straightforwardly extended to include judge 3, judge 4, and so on—each with their own private preference. Without an incentive to coordinate their decisions (e.g., if every $C_i = 0$), judges will heed their personal preference—resulting in interpretive disagreement (i.e., the top-right outcome in Table 1).

Yet the legitimacy of legal (and some nonlegal) systems depends on their stability and predictability (23, 24): comparable cases, which may occur at separate moments in time, should be decided consistently, even by different judges. For a legal system to exhibit stability and predictability, judges must be rewarded for coordinating their interpretations of legally comparable cases (25). If judges' payoff from coordination is greater than the payoff received by satisfying their private preferences (i.e., $C_1 > P_1$ and $C_2 > P_2$), they will seek to choose among the potential equilibria (i.e., the top-left or bottom-right outcomes).

How might individual judges contribute to a legal system's expression of stability? In some cases, judges may consult records of past decisions or deliberate and seek consensus with their peers. Our present research uncovers a further means through which legal officials coordinate their interpretations by default: even without communicating, judges can achieve coordination by treating the rules' text as a default coordination device or focal point (20, 25)—coordinating around conviction (bottom-right) in overinclusion cases and acquittal (top-left) in underinclusion cases.

This focal point theory of statutory interpretation was supported by multiple strands of evidence: (1) whereas laypeople demonstrated substantial variability within and across cultures, legal experts achieved greater interpretive agreement—and did so by adhering to the rule's literal meaning. (2) When offered monetary incentives to coordinate their interpretations with an anonymous partner (whose private preferences would be unknown), laypeople acquired stronger textualist tendencies than when individually judging the same set of cases. (3) Our results revealed a common mechanism underlying the effects of legal expertise and coordination: laypeople's statutory interpretation was guided by their personal moral preferences (i.e., their attitudes of moral blame), while moral preferences had no effect on lawyers' interpretive judgments or on lay participants offered coordination incentives.

The present article reports the findings of a large-scale survey experiment on statutory interpretation conducted in 15 countries. The studies employed a series of vignette pairs, with an overinclusion and an underinclusion case in each pair. Each vignette described an incident (e.g., a fatal traffic accident involving an inebriated driver), followed by a description of the rule or law to which it gave rise. Thereafter, the vignette described a target act, either an overinclusion case (e.g., driving after using alcohol-based mouthwash) or an underinclusion case (e.g., driving after using ecstasy).

In our primary study, participants were randomly assigned to one of 12 conditions in a 2 (case type) × 3 (scenario) × 2 (evaluation mode) between-subjects design. Each participant considered one of three rules and evaluated one case (overinclusion or underinclusion) in either the separate or joint evaluation mode (see *Materials and Methods*). In every condition, participants judged whether the protagonist had violated the rule (the primary transgression judgment). Participants in the joint evaluation mode were asked two additional questions: (i) whether the rule's literal meaning proscribed the target act (e.g., whether the driver ingested alcohol) and (ii) their moral attitude toward the case (i.e., whether the driver's behavior was morally blameworthy).

Results

Our first sequence of analyses examined responses from 4,120 lay participants recruited throughout 15 countries (mean n/country = 275; see Table 2). To ascertain whether our case-type manipulation was effective, we assessed the effect of case type (overinclusion vs. underinclusion) on participants' auxiliary judgments of literal meaning and moral blame in the joint evaluation mode: as expected, overinclusive cases were seen as proscribed by the literal meaning more than underinclusive

Table 1. Statutory interpretation as a mixed-motive coordination game

		Judge 2 at time 2	
		Acquit 2	Convict 2
Judge 1 at time 1	Acquit 1	C_2 / $P_1 + C_1$	P_2 / P_1
	Convict 1		$P_2 + C_2$ / C_1

Note. Individuals' moral values and interpretive commitments (i.e., to textualism vs. purposivism) can engender conflicting private preferences, P_i, for one verdict over another. In the present example, judge 1 prefers to acquit the agent and receives a payoff for acquittal (P_1) and judge 2 prefers to convict the agent and receives a payoff for conviction (P_2). When coordination is not rewarded (i.e., $C_i = 0$) or weakly rewarded ($P_i > C_i$), judges act on their private preferences and their verdicts manifest interpretive disagreement (i.e., acquit 1 and convict 2). When coordination is strongly rewarded (i.e., $C_i > P_i$), judges seek an equilibrium strategy (i.e., acquit 1 and convict 1 or acquit 2 and convict 2). The text of a statute—due to its salience and/or greater univocality—operates as a focal point in such circumstances, facilitating coordination among multiple judges in the absence of communication. Rewards on coordination can arise from formalist sources (e.g., commitment to a legal system's stability and predictability) or realist sources (e.g., one's reputation and career advancement).

Table 2. Sample composition

Country	N	Age mean (SD)	Gender (% women)	Recruitment method
Brazil	207	27.1 (9.83)	52%	Word-of-mouth
Canada	206	34.7 (12.0)	48%	Panel (www.prolific.co)
Colombia	259	22.0 (3.80)	35%	Extra credit
Finland	142	30.3 (13.4)	40%	Panel
Germany	359	37.0 (11.4)	50%	Panel (www.clickworker.de)
India	254	32.5 (9.91)	37%	Panel (www.qualtrics.com)
Italy	319	30.4 (10.9)	23%	Panel (www.prolific.co)
Latvia	569	37.8 (10.4)	63%	Panel (www.qualtrics.com)
Lithuania	191	32.8 (9.18)	39%	Word-of-mouth
Mexico	210	24.4 (5.04)	39%	Panel (www.prolific.co)
Netherlands	391	45.6 (16.7)	45%	Panel (www.panelinzicht.nl)
Poland	271	29.0 (8.61)	43%	Word-of-mouth
Spain	286	43.2 (15.3)	55%	Panel (www.netquest.com)
United Kingdom	202	33.6 (12.7)	70%	Panel (www.prolific.co)
United States	254	37.4 (11.2)	48%	Panel (www.mturk.com)
Total	4120	36.0 (14.1)	46%	—

cases ($B = 2.26$, $t = 25.00$, $\eta^2 = 0.23$), while underinclusive cases were seen as more morally blameworthy than overinclusive cases ($B = -3.24$, $t = -41.70$, $\eta^2 = 0.46$; both Ps < 0.001).

Turning to our primary analysis, a mixed-effects model of transgression judgments revealed an effect of case type, which was qualified by the two-way interaction with evaluation mode (Table 3 and *SI Appendix*, Analysis 1). Replicating previous evidence (14), the effect of case type indicated that overinclusive cases ($M = 4.23$) were more likely to be considered transgressions than were underinclusive cases ($M = 3.73$; $B = 0.51$, $t = 7.23$, $\eta_p^2 = 0.01$, $P < 0.001$; see also Fig. 1*A*).

In addition to judging whether the agent had violated the rule, participants in the joint evaluation mode reported whether the rule's literal meaning proscribed the act and whether the act was morally blameworthy. This provided the opportunity to conceptually replicate our primary finding by regressing transgression judgments on ratings of literal meaning and moral blame. In this model, literal meaning ($B = 0.54$, $t = 29.22$, $\eta_p^2 = 0.30$) and moral blame ($B = 0.15$, $t = 8.53$, $\eta_p^2 = 0.03$) independently predicted transgression judgments (both Ps < 0.001). In sum, laypeople's approaches to statutory interpretation throughout 15 countries reflected both textual and moral criteria—though the influence of the former appeared to be substantially stronger overall (13, 14).

Examining Cultural Variation. An aggregate tendency toward textualism could mask the presence of variability across cultures. Treating country as a fixed factor in the primary regression model uncovered substantial variation in transgression judgments across countries, as indicated by the country × case type interaction (see Table 3). The simple effect of case type

revealed a tendency toward textualist interpretation in Brazil ($P = 0.021$), Canada ($P = 0.013$), Finland, Germany, Italy, Lithuania, and Poland (Ps < 0.001). The effect of case type was nonsignificant in Mexico ($P = 0.081$), Colombia, India, Latvia, the United Kingdom and the United States (Ps > 0.11)—and reversed in two countries, namely, Spain ($P = 0.002$) and the Netherlands ($P = 0.010$).

We defined each country's textualism score as the marginal effect of case type (across rules and evaluation modes)—with positive values representing greater transgression judgments in overinclusion cases than underinclusion cases. Fig. 1*B* displays textualism scores for each country.

To understand whether cultural differences in statutory interpretation were tied to variability in the effects of moral blame and/or literal meaning, we devised an additional test with country ($k = 15$) as the unit of analysis. We treated the by-country regression coefficients of moral blame and literal meaning (drawn from the joint evaluation mode) as indicators of cultural emphases on moral and textual standards, respectively. We then correlated these measures with textualism scores obtained from an independent sample drawn from the same country (i.e., responses in the separate evaluation mode).

In Fig. 1*C*, we plot the regression coefficients of literal meaning and moral blame (on the x axis) against textualism scores (on the y axis). The effect of literal meaning did not predict textualism at the national level (Spearman's $\rho = 0.08$, $P = 0.79$), whereas the effect of moral blame did (Spearman's $\rho = -0.55$, $P = 0.036$). In other words, cultural differences in statutory interpretation were explained by variability in the extent to which moral blame influenced transgression judgments. Including the legal expert data in these analyses ($k = 19$) confirmed

Table 3. Mixed-effects models of transgression judgments

		Laypeople				Legal experts			
		F	dfs	p	η_p^2	F	dfs	p	η_p^2
Preregistered model	Case type	52.98	(1, 4106)	< 0.001	0.013	97.84	(1, 766)	< 0.001	0.113
	Evaluation mode	1.93	(1, 4103)	0.16	0.001	4.59	(1, 766)	0.032	0.006
	Case type × eval. mode	16.52	(1, 4106)	< 0.001	0.004	2.61	(1, 767)	0.11	0.003
Exploratory model	Country	5.11	(14, 4086)	< 0.001	0.018	4.02	(3, 763)	0.007	0.015
	Case type × country	9.42	(14, 4086)	< 0.001	0.031	7.57	(3, 763)	< 0.001	0.029

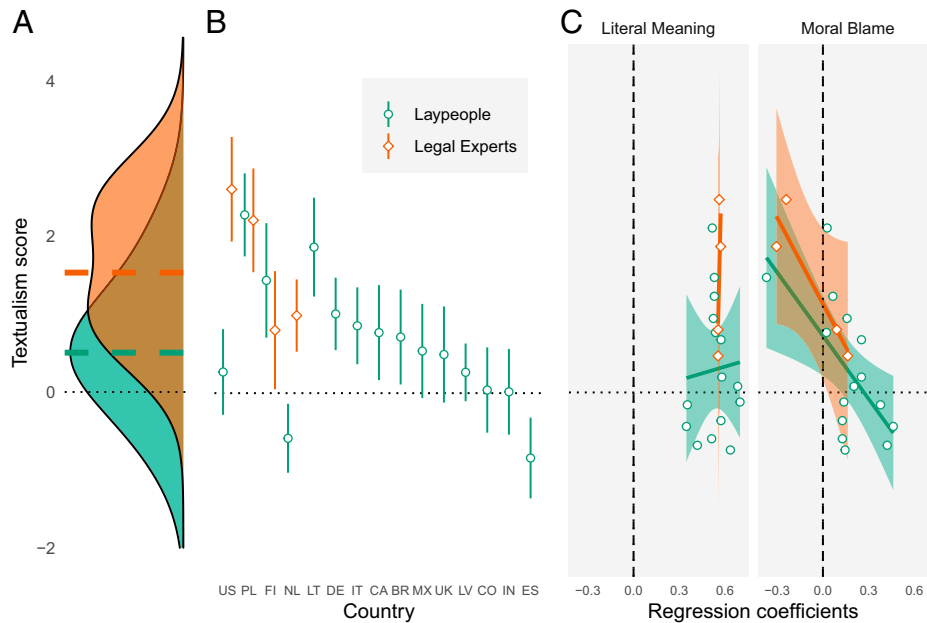Note. Degrees of freedom (*df*s) are calculated using the Kenward–Roger approximation.

Fig. 1. Textualism scores among laypeople and legal experts: *A–C* share a common *y* axis that displays textualism scores. Positive textualism scores represent the tendency to treat overinclusion cases as greater transgressions than underinclusion cases. Negative scores represent the tendency to treat underinclusion cases as greater transgressions than overinclusion cases. (*A*) Grouped density plot by expertise (laypeople vs. legal experts) and overlaid group means. (*B*) National textualism scores and 95% CIs. Countries are placed along the *x* axis, using two-letter country codes: US = United States, PL = Poland, FI = Finland, NL = The Netherlands, LT = Lithuania, DE = Germany, IT = Italy, CA = Canada, BR = Brazil, MX = Mexico, UK = United Kingdom, LV = Latvia, CO = Colombia, IN = India, and ES = Spain. (*C*) National textualism scores in the separate evaluation mode against the regression coefficients of literal meaning and moral blame in the joint evaluation mode. The *x* axes plot the multiple regression coefficients obtained by regressing transgression judgments simultaneously on literal meaning and moral blame ratings—separately for each country. Positive values represent an independent, positive effect of literal meaning (*Left*) or moral blame (*Right*) on transgression judgments—according to the multiple regression model. A value of zero on the *x* axis implies the absence of an effect of the predictor on transgression judgments.

that differences in the coefficient of moral blame predicted variation in textualism scores (Spearman's $\rho = -0.69$, $P = 0.002$), whereas differences in the coefficient of literal meaning did not (Spearman's $\rho = 0.02$, $P = 0.92$).

This evidence hints toward the influence of sociocultural factors and legal traditions in shaping statutory interpretation. To explore these relationships, we conducted further by-country correlation analyses (*SI Appendix*, Analysis 2) but did not find that statutory interpretation differed between common and civil law traditions, countries with a stronger versus weaker adherence to the rule of law, or along cultural and economic dimensions.

Elevated Textualism among Legal Experts. As part of our main study, we also recruited 775 legal experts (596 legal professionals and 197 law students) from four countries: Finland, the Netherlands, Poland, and the United States (mean *n*/country = 194). Manipulation checks confirmed that legal experts perceived (i) overinclusive cases as proscribed by the rule's literal meaning to a greater extent than underinclusive cases ($B = 2.39$, $t = 11.74$, $\eta^2 = 0.27$) and (ii) underinclusive cases as more morally blameworthy than overinclusive cases ($B = -3.52$, $t = -20.29$, $\eta^2 = 0.52$; both $Ps < 0.001$).

Our primary analysis uncovered a large effect of case type (overinclusion vs. underinclusion) and a small effect of evaluation mode. This time, the two-way interaction was not statistically significant (Table 3 and *SI Appendix*, Analysis 1). The main effect of case type indicated that overinclusion cases ($M = 4.67$) were more likely to be considered transgressions than underinclusion cases ($M = 3.14$; $B = 1.53$, $t = 9.91$, $\eta_p^2 = 0.11$, $P < 0.001$)—a pattern that arose in all four countries when analyzed separately (Finland $P = 0.037$, remaining $Ps < 0.001$).

To evaluate the effect of legal expertise, we compared lawyers' and law students' judgments with those of laypeople

drawn from the same four countries, employing propensity score matching (26, 27) to eliminate the imbalance in age, gender, and nationality between lay and expert groups (*SI Appendix*, Analysis 3). We matched ($n_{pairs} = 758$) participants in the experimental (i.e., expert) group to their "nearest neighbor" in the control group based on their predicted probability of being legal experts (i.e., their propensity scores)—thereby reducing covariate imbalance between the lay and expert samples. In this matched dataset, we ran a mixed-effects model entering the expertise term and observed an expertise × case type interaction ($F = 23.18$, $\eta_p^2 = 0.02$, $P < 0.001$). The simple effects of expertise indicated that legal experts were less likely than the matched group of laypeople to view underinclusive cases as transgressions ($B = -0.69$, $t = -4.43$, $P < 0.001$) and more likely to judge overinclusive cases as transgressions ($B = 0.40$, $t = 2.47$, $P = 0.014$) (see Fig. 2*B*).

Moderation analyses in the joint evaluation condition revealed no main effect of expertise ($F = 0.40$, $P = 0.53$) or expertise × literal meaning interaction ($F = 0.90$, $P = 0.34$). An expertise × moral blame interaction did emerge ($F = 10.13$, $\eta_p^2 = 0.01$, $P = 0.002$). Specifically, moral blame predicted transgression judgments among laypeople ($B = 0.12$, $t = 2.92$, $P = 0.004$), but not legal experts ($B = -0.06$, $t = -1.51$, $P = 0.13$; see Fig. 2*A*). *SI Appendix*, Analysis 3 reveals qualitatively indistinguishable results when comparing legal professionals and law students with the entire (unmatched) lay sample.

In sum, legal experts revealed stronger textualist tendencies than did laypeople. When issuing transgression judgments, experts appeared to consider solely the rule's literal meaning, while disregarding their moral preferences. Thus, the discrepancy between experts and laypeople arose partly due to the influence of moral blame on transgression judgments among the latter, but not among the former.
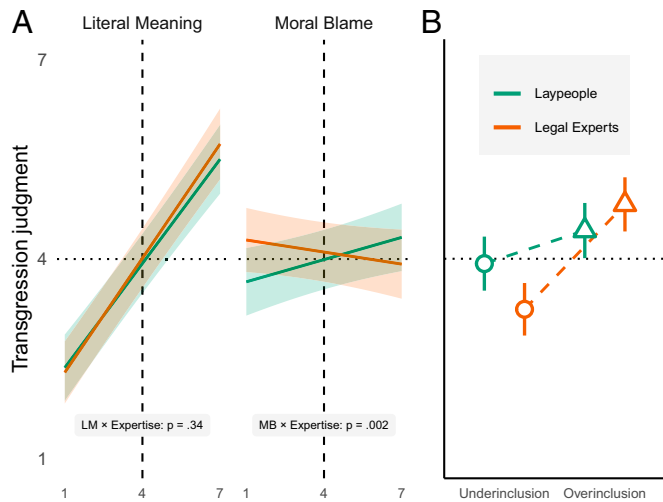
Fig. 2. Expertise effect on transgression judgments. *A* and *B* share a common *y* axis that displays transgression judgments on a seven-point Likert scale. Higher values represent greater agreement with a statement that the agent violated the rule (1 = "strongly disagree," 7 = "strongly agree"). (*A*) Conditional effect plots of literal meaning (*Left*) and moral blame (*Right*) by expertise (laypeople vs. legal experts). The *x* axes span the scale range of literal meaning and moral blame ratings, with higher values reflecting agreement with statements that the agent violated the literal meaning of the rule (*Left*) and that their conduct was morally blameworthy (*Right*). The moral blame × expertise interaction was statistically significant (*P* = 0.002), whereas the literal meaning × expertise interaction was not (*P* = 0.34). LM = literal meaning; MB = moral blame. (*B*) Mean transgression judgments and 95% CIs by case type and expertise (laypeople vs. legal experts). Case type is placed on the *x* axis, with underinclusive cases on the *Left* (circles) and overinclusive cases on the *Right* (triangles).

Text as Focal Point in a Coordination Game. Finally, we explored whether incentives on coordination underlie the tendency toward textualism in interpretive contexts. Our empirical prediction builds on the recognition that statutory interpretation is governed by a norm rewarding predictability and consistency across cases. We hypothesize that these norms of legal decision making instill in legal experts an incentive to coordinate their interpretations, and in these circumstances, the rule's literal meaning—and not its purpose—acts as a focal point (20, 21).

To evaluate this prediction, we examined people's interpretive judgments in an incentivized, two-player coordination game. In the control condition, participants were asked to issue transgression judgments for a series of eight cases. Meanwhile, in the coordination condition, participants were randomly paired with an anonymous partner and each player was offered a monetary reward for matching their transgression judgments with their partner without communicating. If a rule's literal meaning serves as a focal point, the incentive to coordinate should strengthen participants' reliance on literal meaning. We analyzed the data in a mixed-effects logistic regression with case type (overinclusion vs. underinclusion), condition (control vs. coordination), and the case type × condition interaction as fixed effects (treating participants and scenarios as crossed random effects). This model revealed an effect of case type ($\chi^2 = 135.35$) and a case type × condition interaction ($\chi^2 = 24.28$, both *P*s < 0.001). No main effect of condition was observed ($\chi^2 = 0.29$, *P* = 0.59).

As predicted, the case type × condition interaction indicated that (i) overinclusion cases were more likely to be considered transgressions in the coordination condition (prob. = 0.62) than in the control condition (prob. = 0.53; odds ratio [OR] = 1.48, *z* = 3.84, *P* < 0.001) and (ii) underinclusion cases were less likely to be considered transgressions in the coordination

condition (prob. = 0.32) than in the control condition (prob. = 0.39; OR = 0.72, *z* = −3.15, *P* = 0.002; see Fig. 3*B*)—unveiling stronger textualist tendencies under conditions promoting coordinated interpretation.

To ascertain whether coordination incentives strengthened textualist interpretation by reducing participants' emphasis on moral blame (as in the comparison between experts and laypeople), an additional sample (*n* = 299) was asked to provide literal meaning and moral blame ratings for each of the cases. We then calculated mean literal meaning and moral blame ratings for each case and entered these values as case-level predictors in a mixed-effects logistic model of transgression decisions. The model included literal meaning, moral blame, condition (control vs. coordination), and the literal meaning × condition and moral blame × condition interactions as fixed effects. This analysis revealed a main effect of literal meaning ($\chi^2 = 57.56$, *P* < 0.001) and both literal meaning × condition ($\chi^2 = 4.33$, *P* = 0.037) and moral blame × condition interactions ($\chi^2 = 7.64$, *P* = 0.006). No main effects of condition or moral blame were observed (*P*s > 0.16). Whereas literal meaning predicted transgression decisions in both control (*z* = 7.00, OR = 2.20) and coordination (*z* = 8.21, OR = 2.96) conditions (both *P*s < 0.001), the effect of moral blame was significant in the control (*z* = 2.43, OR = 1.49, *P* = 0.015), but not the coordination (*z* = −0.64, OR = 0.89, *P* = 0.52), condition (see Fig. 3*A*). In sum, when experimentally incentivized to coordinate their interpretive judgments, participants tended to disregard their moral preferences and strengthen their adherence to the rules' literal meaning—as stipulated by the focal point theory of statutory interpretation (see Table 1). As such, these results point toward a common mechanism underlying the effects of legal expertise and experimentally induced coordination on textualist interpretation.
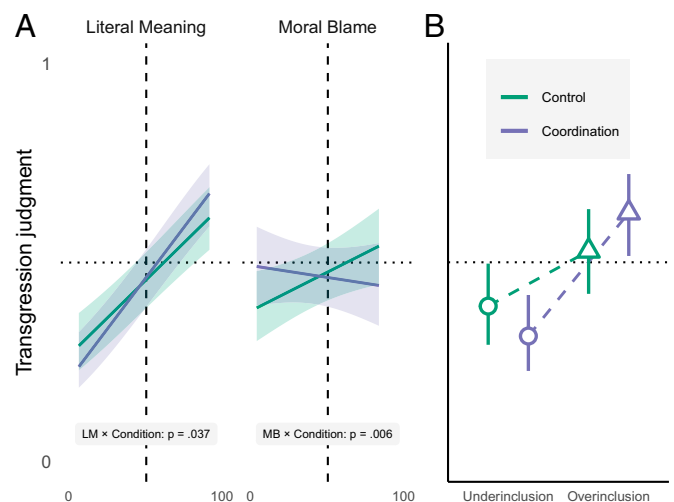

Fig. 3. Coordination effect on transgression judgments. *A* and *B* share a common *y* axis that displays the predicted probability of a transgression judgment. Higher values represent a greater probability of affirming that the agent violated the rule (1 = "yes," 0 = "no"). (*A*) Conditional effect plots of case-level literal meaning (*Left*) and moral blame (*Right*) by condition (control vs. coordination). As in Fig. 2*A*, the *x* axes span the scale range of literal meaning and moral blame ratings, with higher values reflecting agreement with statements that the agent violated the literal meaning of the rule (*Left*) and that their conduct was morally blameworthy (*Right*). Condition interacted with both literal meaning (*P* = 0.037) and moral blame (*P* = 0.006), such that literal meaning had a stronger effect and moral blame had a weaker effect in the coordination condition (relative to the control condition). (*B*) Mean transgression judgments and 95% CIs by case type and condition. Case type is placed on the *x* axis, with underinclusive cases on the *Left* (circles) and overinclusive cases on the *Right* (triangles).

Discussion

A cross-cultural survey experiment documented substantial variability in statutory interpretation across 15 diverse cultures and jurisdictions. Legal experts and laypeople recognized that underinclusive acts (e.g., driving after taking ecstasy) are morally blameworthy, whereas overinclusive acts (e.g., driving after using alcohol-based mouthwash) are not. Nevertheless, when reasoning about which acts violated the law (e.g., a zero-tolerance policy), in the aggregate, participants tended to reach the opposite conclusion: namely, that underinclusive acts comply with the corresponding rules, while overinclusive acts violate them—demonstrating a textualist response pattern. This tendency to prioritize a rule's literal interpretation was further strengthened by legal expertise.

Why would legal experts especially disregard their moral sense and privilege the letter of the law when tasked with applying written rules? Like laypeople, legal professionals hold varied moral views—agreeing or disagreeing with certain legal rules. Various professional incentives, however, discourage legal experts from moralizing rule interpretation: judges seek to avoid being overruled, and lawyers' ethics requires advising their clients of the likely, not personally favored, outcome. More broadly, the rule of law and the legitimacy of judicial decisions hinge on legal systems' expression of stability and predictability in judicial outcomes. Our studies suggested that this circumstance can be fruitfully modeled as a mixed-motive game in which legal officials—despite their heterogeneous moral preferences—can reach an equilibrium if they are rewarded for their coordination. As evidence in favor of this account, lawyers achieved greater interpretive agreement by applying textual standards, and their elevated textualist tendencies were partly explained by a dissociation between their moral attitudes and their interpretive judgments (see also refs. (28, 29)). Furthermore, we experimentally recreated this phenomenon by monetarily incentivizing lay participants to coordinate their interpretive judgments without communication.

Our studies included various mundane rules (e.g., household or workplace rules), which even nonlawyers would be tasked with enforcing. Evidence that laypeople demonstrate textualist inclinations when judging nonlegal cases points toward the broader applicability of our findings and reveals that textualism is not circumscribed to the legal domain. Rather, textualism may be better explained as emerging from the social dimension of legal and nonlegal rules alike (30), i.e., the tendency for rules to govern the conduct of a diverse group of individuals. Absent this social quality, e.g., in the context of personal rules (*SI Appendix,* Analysis 4), the demand for stability and predictability may be relaxed—rendering purposive interpretation more advantageous.

Previous scholarship has theorized that the "plain meaning of a text as applied to a set of facts" can play the role of a coordination device ((31), p. 1557; see also refs. (20, 25)), a salient element of the context that highlights one among multiple equilibria. Our final experiment vindicated this prediction, demonstrating that—when incentivized to coordinate their interpretations of legal and nonlegal rules in circumstances that preclude communication—people strengthen their adherence to the rules' literal meanings.

Implications and Limitations. These results inform ongoing legal debates about the interpretation of contracts, statutes, and constitutions. For example, in American legal interpretation, modern textualist judges increasingly aim to interpret laws in line with what those laws communicate to an ordinary member of the public (see, e.g., ref. (32)). The results here suggest some support for this theory's focus on text: ordinary people's understanding of legal rules is heavily informed by the rules' text.

The findings also reveal a pronounced effect of legal training on the interpretation of rules. Legal experts were more inclined to rely on the letter over the spirit of the law. The coordination game results suggest that legal experts' real-world convergence on literal meaning might not necessarily reflect those experts' consensus about the rule's "ordinary public meaning." The same convergence could also be explained by a rational response to coordination incentives. In other words, experts' real-world coordination around rules' text might reflect their desire to coordinate around a clear focal point.

Our coordination game shares important features of real-world judicial decision making. Commentators note that judges dislike having their decisions reversed on appeal (33) and care deeply about the regard of their peer and popular audiences (34). These interests produce incentives to coordinate (e.g., with appellate judges or with popular reception). However, communicative coordination can be costly or even impossible: judges often manage a large number of cases (35), and there is little time to survey one's peers to identify the outcome on which to coordinate. Moreover, lower court judges who prefer nonreversal would want to know the views of the appellate judge assigned to their case, but in systems that randomly assign judges, the appellate judge's identity—and, by extension, their views—are unknown at the time that the trial court judge evaluates the case. Our economic game, involving incentivization without communication, offers a useful model of this common dynamic and further supports that text serves as a default coordination device in the absence of communication (25, 31).

Though we noted that approaches to statutory interpretation varied substantially across field sites, whether this variation was driven strictly by elements of culture or legal tradition is unclear (*SI Appendix,* Analysis 2). Since our sampling methods differed across locations, variation in the tendency toward textualism could also partially arise from unobserved differences in the samples' composition. Given these sampling differences, we caution readers against drawing strong conclusions about the role of culture or legal tradition in statutory interpretation from our present findings.

Why precisely literal meaning provides a focal point cannot be gleaned from our present studies. One possibility, supported by preliminary data (*SI Appendix,* Analysis 5), is that individuals in diverse communities have a similar understanding of the rule's literal meaning but are prone to disagree in their appraisals of whether the incident violated the rule's deeper purpose. The recognition of greater univocality in literal meaning could instigate coordination around the rule's text over its (morally divisive) purpose. As a future test of this hypothesis, we envision studies of legal reasoning in morally homogeneous societies, in which moral preferences may be more uniform—potentially obviating the need for legal text as a coordination device and helping to establish a link between the emergence of legal text and moral diversity.

In these studies, our focus was on difficult cases involving conflict between literal meaning and moral attitudes. Meanwhile, most real-world incidents simultaneously violate (i.e., true positives) or comply with (i.e., true negatives) both the text and the purpose of a rule, so naturally occurring instances of overinclusion and underinclusion are likely to be infrequent (36). This approach places certain limits on the ecological validity of our findings but in turn offers critical insight into the cognitive basis of legal reasoning by dissociating the roles of the letter versus the spirit of the law.

Finally, our interest in this work was in whether a behavior violates a given rule—what we called transgression judgments. This question is distinct from questions of whether the behavior warrants punishment and, if so, of what magnitude. On the basis of past research (37), it stands to reason that punishment allocations may recruit distinct cognitive processes and reflect a different balance of textual and moral appraisal than was observed when examining transgression judgments.

Conclusions. As part of normative development, adults abandon the uncritical deference to the rule of authority in order to manifest deeper ethical principles (3). Yet when prompted to decide which behaviors are permissible by legal standards, people disregard their personal moral values to a surprising degree and prioritize the literal meaning of rules instead. This textualist approach to interpretation is strengthened by legal training, and evidence from an incentivized experiment yielded potential insight into its origin: applying a rule's literal meaning, in detriment of its intended purpose or instrumental value, can serve as a focal point (20, 25) among individuals who share an interest in aligning their interpretations. In this way, adopting a textualist policy—even while incurring moral costs in certain, rare instances—could facilitate long-term social coordination (38) among lawmakers, citizens, and judges.

Materials and Methods

The studies were conducted with approval from Yale's Human Research Protection Program. Participants were informed about the nature of the study and asked to provide written consent before taking part in the study. Study data, analysis scripts, and stimuli (including translations) are publicly accessible on the *Open Science Framework* at https://osf.io/yw8ek/.

Materials. Our studies employed a battery of nine vignette pairs with one overinclusion and one underinclusion case in each pair. The coordination game made use of eight vignette pairs (vehicles, sleep, driving, library, classroom, shoes, environment, and music), while the main study employed three pairs (classroom, phone, and driving).

The vignettes first described an incident (e.g., "A 21-year-old woman suffered a traffic accident that took her life. The young woman was driving under the influence."), followed by a description of the rule or law to which it gave rise, including its underlying purpose ("In order to avoid future accidents, Congress passed a zero-tolerance policy establishing that: 'If the breathalyzer detects any trace of alcohol, the vehicle will be seized and the driver subject to imprisonment.'"). Then, the vignette described a target act, either in violation of the text of the rule, but not its underlying purpose (in overinclusion cases, e.g., using alcohol-based mouthwash prior to driving), or in violation of the purpose of the rule, but not its text (in underinclusion cases, e.g., using ecstasy prior to driving).

Measures.
Transgression judgment. Our dependent measure was whether the protagonist who carried out the target act had violated the rule. In the main study, transgression judgments (e.g., "Andrea violated the zero-tolerance policy.") were made on a seven-point scale ranging from 1: "strongly disagree" to 7: "strongly agree." In the coordination game, transgression judgments ("Did [the agent] break the rule?") were dichotomous: 1 = "Yes" and 0 = "No."
Supplementary ratings: literal meaning and moral blame. In the main study, participants in the joint evaluation mode were also asked to rate whether the text of the rule proscribed the target act (literal meaning, e.g., "Andrea drove after ingesting a product containing alcohol.") and whether the protagonist's behavior was morally blameworthy (moral blame, e.g., "Andrea is morally blameworthy for what she did."). Both assessments, i.e., of literal meaning and moral blame, were made on seven-point scales ranging from 1: "definitely not" to 7 "definitely."

In the addendum to the coordination game, participants were asked to rate whether the text of the rule proscribed the target act (literal meaning, e.g., "John wore shoes in the house.") and whether the protagonist's behavior was morally blameworthy (moral blame, e.g., "What John did was morally wrong."). Both assessments were made on sliding scales ranging from 0: "strongly disagree" to 100: "strongly agree."

Textualism score. The marginal effect of case type (with underinclusion as the reference level and averaged over levels of evaluation mode and rule) constituted our by-country measure of textualism ($M = 0.89$, $SD = 0.92$). Textualism scores were normally distributed (Shapiro–Wilk test: $W = 0.96$, $P = 0.64$) and strongly correlated across evaluation modes ($r = 0.70$, $P < 0.001$).

Participants.
Laypeople. Four thousand one hundred and twenty participants were recruited in 15 countries (see Table 2 for demographic information and recruitment details).
Legal experts. Five hundred ninety-six law graduates and 179 law students (age: $M = 40.5$, $SD = 13.9$; 48% women) were recruited from four countries: Finland ($n = 124$; 110 law graduates and 14 law students), the Netherlands ($n = 331$; 331 law graduates and no law students), Poland ($n = 161$; 145 law graduates and 16 law students), and the United States ($n = 159$; 9 law graduates and 150 law students).
Coordination game. Six hundred participants (age: $M = 26.4$, $SD = 8.61$; 40% women) were recruited via Prolific.co and invited to take part in an experiment in exchange for monetary compensation.
Coordination game: addendum. Two hundred ninety-nine participants (age: $M = 37.6$, $SD = 12.0$; 49% women) were recruited via Prolific.co and invited to take part in an experiment in exchange for monetary compensation.

Procedure: Main Study. In a 2 (case: overinclusive and underinclusive) × 2 (evaluation mode: separate and joint) × 3 (scenario: car, phone, and alcohol) between-subjects design, participants read either an overinclusion or an underinclusion case.

Our primary dependent measure was participants' agreement or disagreement with a statement that the agent had violated the rule. In the joint evaluation mode, the primary dependent measure was accompanied by two supplementary assessments of the literal meaning of the rule and the agent's moral blame (see *Measures* subsection).

Procedure: Coordination Game. In a 2 between- (condition: control and coordination) × 2 within- (case: overinclusive and underinclusive) × 8 within- (scenario) balanced incomplete block design, participants read a sequence of six scenarios (plus two filler trials). In the control condition, participants were asked to "make a decision: did the person violate the rule (YES) or not (NO)?" Meanwhile, in the coordination condition, participants were told:

> "You are invited to play the Judging Game. You are Judge 1 and you have been paired with another player, Judge 2. On the following screens, both of you will be reading the same eight stories. Each story describes a rule and a person's behavior. After reading each story, you will both be asked to make a decision: Did the person violate the rule (YES) or not (NO)?
>
> To win extra earnings, you and Judge 2 must agree on as many decisions as possible. You must try and reach the same decision on Case 1, on Case 2, on Case 3, etc., all the way through Case 8 without talking to each other. If you agree on at least six decisions, each of you will earn an additional £1.00 (for a total of £1.70). If not, neither of you will earn the additional £1.00."

Participants made a dichotomous transgression judgment for each scenario. At the end of the study, participants in the coordination condition were randomly paired and paid a £1 bonus if they agreed on at least six of the eight cases. Study design, predictions, and analysis plans were preregistered at https://aspredicted.org/qj5mc.pdf.

Procedure: Coordination Game Addendum. In a 2 within- (case: overinclusive and underinclusive) × 8 within- (scenario) balanced incomplete block design, participants read a sequence of six scenarios (plus two filler trials). After each case, participants were asked to judge whether the case violated the literal meaning of the rule and whether the agent was morally blameworthy (see *Measures*).

Data, Materials, and Software Availability. Anonymized study data, analysis scripts, and stimuli (including translations) have been deposited in the Open Science Framework (https://osf.io/yw8ek/) (39).

Author affiliations: ªUniversidad de Granada, 18071 Granada, Spain; ᵇGeorgetown University, Washington, DC 20057; ᶜYale University, New Haven, CT 06520; ᵈPontifical Catholic University of Rio de Janeiro, 22541 Rio de Janeiro, Brazil; ᵉUniversity of Zurich, 8006 Zürich, Switzerland; ᶠJagiellonian University in Kraków, 31007 Kraków, Poland; ᵍUniversiteit Leiden, 2311 Leiden, the Netherlands; ʰVilnius University, 01513 Vilnius, Lithuania; ⁱUniversity of Silesia in Katowice, 40007 Katowice, Poland; ʲUniversity of Helsinki, 00100 Helsinki, Finland; ᵏUniversity College London, London WC1E 6BT, United Kingdom; ˡRīga Stradiņš University, 1007 Riga, Latvia; and ᵐUniversidad Nacional de Colombia, 500001 Bogotá, Colombia

1. J. C. Fell, M. Scherer, S. Thomas, R. B. Voas, Assessing the impact of twenty underage drinking laws. *J. Stud. Alcohol Drugs* **77**, 249–260 (2016).
2. L. P. Nucci, E. Turiel, Social interactions and the development of social concepts in preschool children. *Child Dev.* **49**, 400–407 (1978).
3. L. Kohlberg, *The Philosophy of Moral Development*, Essays on Moral Development (Harper & Row, 1981), vol. **I**.
4. C. S. Sripada, S. Stich, "A framework for the psychology of norms" in *The Innate Mind: Volume 2: Culture and Cognition*, P. Carruthers, S. Laurence, S. Stich, Eds. (Oxford University Press, 2005), pp. 280–301.
5. F. Cushman, Action, outcome, and value: A dual-system framework for morality. *Pers. Soc. Psychol. Rev.* **17**, 273–292 (2013).
6. R. M. Miller, I. A. Hannikainen, F. A. Cushman, Bad actions or bad outcomes? Differentiating affective contributions to the moral condemnation of harm. *Emotion* **14**, 573–587 (2014).
7. L. Fuller, Positivism and fidelity to law: A reply to professor Hart. *Harv. Law Rev.* **71**, 630–672 (1958).
8. A. Barak, *Purposive Interpretation in Law* (Princeton University Press, 2005).
9. H. L. A. Hart, Positivism and the separation of law and morals. *Harv. Law Rev.* **71**, 593–629 (1958).
10. F. Schauer, Formalism. *Yale Law J.* **97**, 509–548 (1988).
11. G. L. Priest, B. Klein, The selection of disputes for litigation. *J. Legal Stud.* **13**, 1–55 (1984).
12. J. Bregant, I. Wellbery, A. Shaw, Crime but not punishment? Children are more lenient toward rule-breaking when the "spirit of the law" is unbroken. *J. Exp. Child Psychol.* **178**, 266–282 (2019).
13. S. M. Garcia, P. Chen, M. T. Gordon, The letter versus the spirit of the law: A lay perspective on culpability. *Judgm. Decis. Mak.* **9**, 479–490 (2014).
14. N. Struchiner, I. R. Hannikainen, G. Almeida, An experimental guide to vehicles in the park. *Judgm. Decis. Mak.* **15**, 312–329 (2020).
15. E. Martínez, K. Tobia, What do law professors believe about law and the legal academy? An empirical inquiry. SSRN [Preprint] (2022). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4182521 (Accessed 30 August 2022).
16. J. F. Manning, Textualism and legislative intent. *Va. Law Rev.* **91**, 419–450 (2005).
17. A. S. Krishnakumar, Cracking the whole code rule. *New York Univ. Law Rev.* **96**, 76–172 (2021).
18. C. R. Sunstein, Moral heuristics. *Behav. Brain Sci.* **28**, 531–542, discussion 542–573 (2005).
19. S. Mousavi, G. Gigerenzer, Heuristics are tools for uncertainty. *Homo Oeconomicus* **34**, 361–379 (2017).
20. T. C. Schelling, *The Strategy of Conflict* (Harvard University Press, 1960).
21. R. H. McAdams, A focal point theory of expressive law. *Va. Law Rev.* **86**, 1649–1729 (2000).
22. L. Epstein, C. M. Parker, J. A. Segal, Do justices defend the speech they hate? An analysis of in-group bias on the US supreme court. *Journal of Law and Courts* **6**, 237–262 (2018).
23. L. Fuller, *The Morality of Law* (Yale University Press, 1964).
24. I. R. Hannikainen *et al.*, Are there cross-cultural legal principles? Modal reasoning uncovers procedural constraints on law. *Cogn. Sci. (Hauppauge)* **45**, e13024 (2021).
25. F. Schauer, Statutory construction and the coordinating function of plain meaning. *Supreme Court Rev.* **1990**, 231–256 (1990).
26. P. R. Rosenbaum, D. B. Rubin, The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**, 41–55 (1983).
27. D. E. Ho, K. Imai, G. King, E. A. Stuart, Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Polit. Anal.* **15**, 199–236 (2007).
28. D. M. Kahan *et al.*, Ideology or situation sense: an experimental investigation of motivated reasoning and professional judgment. *Univ. Pa. Law Rev.* **164**, 349–439 (2015).
29. K. P. Tobia, Legal concepts and legal expertise. SSRN [Preprint] (2020). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3536564 (Accessed 30 August 2022).
30. C. Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms* (Cambridge University Press, 2005).
31. W. Eskridge, Textualism, the unknown ideal? *Mich. Law Rev.* **96**, 1509–1560 (1998).
32. Bostock v. Clayton County, 590 U.S. ___ (2020).
33. R. Posner, *The Federal Courts: Challenge and Reform* (Harvard University Press, 1996).
34. L. Baum, *Judges and Their Audiences: A Perspective on Judicial Behavior* (Princeton University Press, 2006).
35. R. Posner, *The Federal Courts: Crisis and Reform* (Harvard University Press, 1985).
36. T. Eisenberg, Testing the selection effect: A new theoretical framework with empirical tests. *J. Legal Stud.* **19**, 337–358 (1990).
37. F. Cushman, Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition* **108**, 353–380 (2008).
38. W. M. Bennis, D. L. Medin, D. M. Bartels, The costs and benefits of calculation and moral rules. *Perspect. Psychol. Sci.* **5**, 187–202 (2010).
39. I. R. Hannikainen, K. P. Tobia, Data, script and materials for "Coordination and expertise foster legal textualism." Open Science Framework. https://osf.io/yw8ek/. Deposited 4 September 2022.