Are There Cross-Cultural Legal Principles? Modal Reasoning Uncovers Procedural Constraints on Law

Ivar R. Hannikainen*, Kevin P. Tobia, Guilherme da F. C. F. de Almeida, Raff Donelson, Vilius Dranseika, Markus Kneer, Niek Strohmaier, Piotr Bystranowski, Kristina Dolinina, Bartosz Janik, Sothie Keo, Eglė Lauraitytė, Alice Liefgreen, Maciej Próchnicki, Alejandro Rosas & Noel Struchiner

* Correspondence concerning this article should be addressed to: Ivar R. Hannikainen, Department of Philosophy I, Faculty of Psychology, Cartuja Campus, Universidad de Granada (Spain), 18011. Contact: ivar@ugr.es

Word count (excluding Abstract, Tables, References and Appendices): 3933

Abstract

Despite pervasive variation in the content of laws, legal theorists and anthropologists have often argued that laws share certain abstract features and even speculated that law may be a human universal. In the present report, we contribute cross-cultural data to this debate: Are there cross-cultural principles of law? Participants in eleven different countries (N = 3054) were asked whether there could be laws that violate certain procedural principles (e.g., laws applied retrospectively or unintelligible laws), and also whether there are any such laws—in a between-subjects design. Confirming our pre-registered prediction, people reported that such laws cannot exist, but also (paradoxically) that there are such laws. These results document cross-culturally and –linguistically robust beliefs about the nature of law which defy people's conception of how legal systems function in practice.

Are There Cross-Cultural Legal Principles? Modal Reasoning Uncovers Procedural Constraints on Law

1. Introduction

Laws vary remarkably from one jurisdiction to the next. Even within jurisdictions, legislative changes are frequent and shift the legal status of various practices over time. These changes in legality accompany fluctuations in the prevailing morality (Ofosu, Chambers, Chen, & Hehman, 2019), perhaps even helping to precipitate shifts in public opinion. Despite the abrupt historical change and cultural diversity in the *content* of legal norms, theorists in law (Fuller, 1964; Finnis, 1980) and anthropology (Brown, 1991) have speculated that certain features of their *form* may be universal. Yet no research to date has examined whether people across cultures share intuitions about what laws fundamentally are. In the present work, we sought to fill this gap in our understanding of the cognitive science of law.

In the neighboring disciplines of moral psychology and behavioral economics, extensive research agendas have now documented robust patterns in people's moral (Cushman, Young, & Hauser, 2006; Rozin, Lowery, Imada, & Haidt, 1999) and economic (Boyer & Pedersen, 2018) preferences across various cultures, including small-scale societies: for instance, the role of intent in moral blameworthiness (Barrett et al., 2016), the distinction between personal and impersonal forms of harm (Awad, Dsouza, Shariff, Rahwan, & Bonnefon, 2020), the tendency to uphold taboos related to bodily purity and sanctity (Graham, Haidt, & Nosek, 2009), such as norms proscribing sibling incest (Haidt, Koller, & Dias, 1993), and systematic deviations from self-interest, such as the tendency to exact costly punishment (Henrich et al., 2005).

Theorists in the legal domain have argued that this palette of moral sentiments forms the basis for an 'intuitive jurisprudence' (Mikhail, 2007; Sznycer & Patrick, 2020), according to which the structure and content of criminal legal codes crystallize basic aspects of our moral sense. In one recent study (Sznycer & Patrick, 2020), participants were asked to

consider a series of ancient laws, drawn from the Tang Code (such as gratuitously killing one's slave) and the Laws of Eshnunna (such as liability for one's goring ox), and reported their moral reactions—e.g., of perceived shame and wrongness—in response to hypothetical violations of these dated laws. Strikingly, the intensity of participants' reactions predicted the magnitude of legal punishment, whether in the form of a fine or prison time, codified in these millennia-old criminal codes. This universalist paradigm—implicating our basic moral sensibilities in the genesis of legal doctrines—would also help explain the near-universal emergence of certain criminal laws (e.g., regarding murder, see Mikhail, 2009) and of legal institutions as a basic property of social groups (Brown, 1991; Hoebel, 1954; Nader, 1965).

In our present study, we pursue a related prediction: namely, that people around the world share an intuitive grasp of the *formal* properties of law—as well as its content. The view that laws must observe certain procedural principles has come to be associated most strongly with the work of American philosopher Lon Fuller. His famous (1964) book told the tale of a hypothetical king, Rex, who—through a sequence of failures—gradually discovered the eight procedural principles capable of transforming his imperatives and royal wishes into what could be properly referred to as a legal system. For instance, at first, Rex did not publicly proclaim the rules of his kingdom, but instead kept them secret in his diary. As a result, the populace could not possibly *know* Rex's rules—which taught Rex his first lesson: that laws need to be publicly promulgated.

Philosophers of law have recurrently debated the question of whether laws must meet certain procedural requirements, but the corresponding body of empirical evidence examining whether the concept of law exhibits such constraints is meager. One recent study—which inspired the present attempt at cross-cultural generalization—did reveal that lawyers and laypeople in the United States consider the procedural principles illustrated in Fuller's (1964) writings to be, in a paradoxical sense, essential to the law (Donelson & Hannikainen, 2020).

The paradox lies in the way participants reacted to different linguistic formulations of the same procedural principles.

The gist of the experimental paradigm is simple: One group of participants was asked to assess whether, in their experience, laws observe or violate each of the procedural principles (e.g., whether or not there are laws *"punishing people for acts that were legal at the time they were performed*"). Looking across all eight procedural principles, participants were divided on whether actual laws abide by these procedural principles.

According to most systems of modal logic, if there exists even a single retrospective law, it follows that retrospective laws certainly *could* exist—in virtue of axiom *M*. A separate group of participants were asked precisely this question: *Could* the laws of a hypothetical nation violate each of the procedural principles (e.g., whether there could be any laws *"punishing people for acts that were legal at the time they were performed*")? Both lawyers and non-lawyers reported the *opposite* pattern of responses: namely, that there could *not* be any retrospective laws (even though there actually *are*). A very similar pattern emerged when considering the remaining procedural principles: e.g., that laws be announced publicly, made intelligible to the vast majority, changed infrequently at most, and so on.

Did this effect arise simply because participants did not grasp the relevant axiom of modal logic (i.e., that if a law exhibits property *P*, then laws necessarily could exhibit property *P*)? A follow-up study spoke against this explanation: When asked to simultaneously consider actual and possible laws, participants reported that laws often abide by the procedural principles, though it would be possible for laws to violate them—reversing the direction of the difference between conditions.

In sum, this preliminary evidence reveals an apparent contradiction: that laws need to observe certain procedural constraints, even though actual laws routinely violate them. One question this raises is: If, from empirical knowledge of what laws are like, people conclude

that there are no procedural constraints that all laws abide by; then how might people form the judgment that laws could not possibly violate these principles?

One possibility is that this judgment stems from a mental representation or concept of 'law' (Margolis & Laurence, 1999). Such a concept could be a prototype, or essentialized belief—which arises from the simpler concepts of 'norm', 'fairness' and 'punishment', for example—and depends weakly on experience with actual laws. This hypothesis yields a further empirical prediction: that the tendency to ascribe these procedural properties to law should emerge across cultures, despite fundamental variation in the particularities of each legal system and its manifestation.

To examine this prediction, we sampled from a diverse set of jurisdictions, including both common (e.g., United States, and India) and civil law (e.g., Brazil, and Poland) countries, while also ascertaining that the jurisdictions vary in the strength of rule of law (see World Justice Project, 2020). Additionally, we administered the study in the local language at each field site, helping to establish whether the effect of modal reasoning on beliefs about the law arises across different language families (from Romance, Germanic, and Balto-Slavic, to Indic and Austroasiatic families).

Finally, this cross-cultural and –linguistic approach enables us to overcome a limitation of previous research in psychological science: Various reports have highlighted the discipline's ethnocentric bias (e.g., Thalmayer, Toscanelli, & Arnett, 2021), i.e., in advancing claims about human psychology on the basis of evidence derived almost exclusively from Western, English-speaking samples.

2. Methods

2.1 Field sites and participants

We selected 11 different field sites for data collection on the basis of (i) linguistic and cultural diversity, and (ii) variation in the strength of the rule of law. To ensure a balance of

strong and weak rule of law, we drew on the Rule of Law Index 2020 (World Justice Project, 2020). The *Rule of Law Index* draws on thousands of household and expert surveys worldwide to quantify the strength of the rule of law across nations, on eight complementary dimensions: constraints on government, absence of corruption and 'revolving doors', open government (guaranteeing information and civic participation), fundamental rights (e.g., absence of discrimination and freedom of expression and assembly), order and security (low levels of crime and political violence), regulatory enforcement, civil justice, and criminal justice. Netherlands and Germany were rated as having very strong rule of law—appearing among the top ten countries. Some other countries (including Brazil, India, Colombia and Cambodia) were classified as having a weak rule of law—and spanned the bottom half of the global ranking.

The minimum target sample size per site was established at 200 participants per field site, and was met everywhere except Cambodia (see Table 1). This sample size provides adequate statistical power ($\beta = .20$) to detect an odds ratio ≥ 1.40 setting α at .05.

Table 1.

Sample characteristics.

Country	***N***	**Age** Mean (SD)	**Gender** (% women)	**Recruitment method**
Brazil	223	27.5 (10.1)	51.0%	Word-of-mouth
Cambodia	100	24.1 (6.31)	55.0%	Word-of-mouth
Colombia	263	22.1 (3.80)	35.4%	Extra-credit
Germany	237	37.1 (11.7)	50.2%	Panel
India	275	32.7 (9.50)	63.3%	Panel (www.qualtrics.com)
Lithuania	242	32.4 (9.35)	43.0%	Word-of-mouth
Netherlands	722	45.9 (14.3)	48.9%	Word-of-mouth & Panel (www.panelinzicht.nl)

Poland	271	29.2 (8.54)	42.3%	Word-of-mouth
Spain	289	43.2 (15.3)	55.1%	Panel (www.netquest.com)
United Kingdom	210	35.2 (12.7)	62.9%	Panel (www.prolific.ac)
United States	222	37.4 (11.4)	57.0%	Panel (www.mturk.com)
TOTAL	3054	36.0 (14.0)	48.5%	-

2.2 Materials

The stimuli were adapted from Donelson & Hannikainen (2020) and translated into eight additional languages by native speakers: Dutch, German, Hindi, Khmer, Lithuanian, Polish, Portuguese, and Spanish. Site collaborators were asked to iteratively compare oral back-translations into English against the original materials.

The main task was made up of eight pairs of statements with an affirmation (e.g., "*Some laws* change very frequently") and a negation (e.g., "*No laws* change very frequently") in each pair. These items could be worded as either existential statements, or modal statements as shown in Figure 1.

Thus, the difference between conditions was the inclusion of an auxiliary verb transforming the existential statements into modal statements. In each language, we sought to employ auxiliary verbs that primarily denote possibility and necessity (i.e., alethic modality; see Table 2).

Table 2.

Modal auxiliary verbs by language.

Language	**auxiliary verb (count)**
English	could (14), would (4), might (1).
Dutch	kan/zouden kunnen (14), moeten/zouden moeten (3), mogelijk (1).

German	könnte/könnten (10), kann (7), können (1).
Hindi	सकना (*sakanā*; 15), होना (*honā*; 3), चाहिए (*cāhiye*; 1).
Khmer	អាច (*ach*; 19), នឹង (*nung*; 4).
Lithuanian	(ne)galėtų (14), turėtų (4), gali (1).
Polish	może/mogą (15), muszą (4).
Portuguese	poderia/poderiam (14), teria/teriam (3).
Spanish	podría/podrían (16), tendría/tendrían (3).

The post-experiment test consisted of three questions, assessing how participants had interpreted the task. On a seven-point scale anchored at -3: "Not at all" and 3: "Completely", reported whether they were thinking about:

(1) "what laws are usually like, in your experience" (i.e., *empirical* interpretation),

(2) "what laws have to be like, as in the requirements for something to count as law" (i.e., *alethic* interpretation)

(3) "what laws should be like, according to your beliefs about right and wrong" (i.e., *deontic* interpretation).

Materials in every language are available on the *Open Science Framework* at: https://osf.io/hn8m5/.

2.3 Procedure

In a between-subjects design, participants were randomly assigned to either the *Actual* or the *Possible* condition, and read the following introduction:

You will be shown eight pairs of statements regarding what laws are like. For each pair of statements, please take a moment to think about which statement better reflects your opinion about what laws are like.

In each condition, the statement pairs were presented in a random order across participants. In the Actual condition, participants read an existential statement and its negation in each pair. Meanwhile, in the Possible condition, each pair contained a modal statement employing an auxiliary verb and its negation. Participants were asked to endorse one statement from each pair. After selecting a statement from each pair, on the following page, participants were asked a set of three questions about their interpretation of the task.

Lastly, participants provided demographic information: age (in years), gender (*Male*, *Female*, *Non-binary*), and legal background (*Law student*, *Legal professional*, *Neither*). Subgroup analyses are reported in Supplementary Analysis 1.

2.4 Predictions and analysis plan

Our primary prediction, sample size determination and analysis plan were pre-registered at aspredicted.org/blind.php?x=ev6nk8. Inspired by previous findings (Donelson & Hannikainen, 2020), we hypothesized greater endorsement of procedural principles in the Possible condition than in the Actual condition.

We coded endorsement as 1 if participants reported that a principle obtains, e.g., selected "The law as enforced *does not* [/*could not*] differ from the law as formally announced". Endorsement was coded as 0 if participants stated that a principle was or could be violated, i.e., selected "The law as enforced *differs* [/*could differ*] from the law as formally announced". In the analyses below, we conducted logistic regression models predicting the probability ($\hat{p}$) that participants endorse the procedural principle(s).

We tested our primary prediction in a mixed-effects logistic regression model entering experimental condition as fixed effect, and participant and principle as crossed random effects. Generalized linear mixed-models were conducted with the *lme4* (Bates, Maechler, Bolker, & Walker, 2015) package. Predicted probabilities (notated as $\hat{p}$) are calculated in the *emmeans* package in *R* version 3.6.2. Data and an accompanying *R* script are available at: https://osf.io/hn8m5/.

3. Results

3.1 Pre-registered analyses

Participants in the Actual condition were more likely to report that laws violated procedural principles (*Actual:* $\hat{p}$ = .54, 95% $CI_{asymptotic}$ [.40, .68]) than participants in the Possible condition were to say that those same principles could be violated (*Possible:* $\hat{p}$ = .81, 95% $CI_{asymptotic}$ [.72, .89]), OR = 3.74, 95% CI [3.41, 4.09], z = 28.43, p < .001.

Treating country as a fixed factor, we observed a main effect of country, $\chi^2(10)$ = 115.0, condition, $\chi^2(1)$ = 823.7, as well as a country×condition interaction, $\chi^2(10)$ = 169.6, all *p*s < .001. The country×condition interaction revealed variation in the magnitude of the effect across cultures, with modest effects emerging in India and the United Kingdom (ORs < 1.60) and large effects in Brazil and Poland (ORs > 7.50). Still, the simple effect of condition was significant in every country when analyzed separately (all *p*s < .010). Similarly, treating principle as a fixed factor revealed a main effect of principle, $\chi^2(7)$ = 1877.4, and a principle×condition interaction, $\chi^2(7)$ = 598.4 (while accounting for the effect of condition, $\chi^2[1]$ = 710.8) all *p*s < .001. The interaction indicated that the magnitude of the effect varied significantly across principles, going from small for the generality principle (OR = 1.79) to large for the consistency principle (OR = 13.97; see also Figure 1). Nonetheless, the simple effect of condition was statistically significant for each of the eight principles individually, all *p*s < .001.

Additionally, our pre-registered prediction emerged robustly across age groups and genders, and among lawyers and non-lawyers alike (as detailed in Supplementary Analysis 1). In sum, generalizing the findings of Donelson & Hannikainen (2020), participants in the Possible condition tended to report that laws could *not* violate various procedural principles even while participants in the Actual condition recognized that they often do.

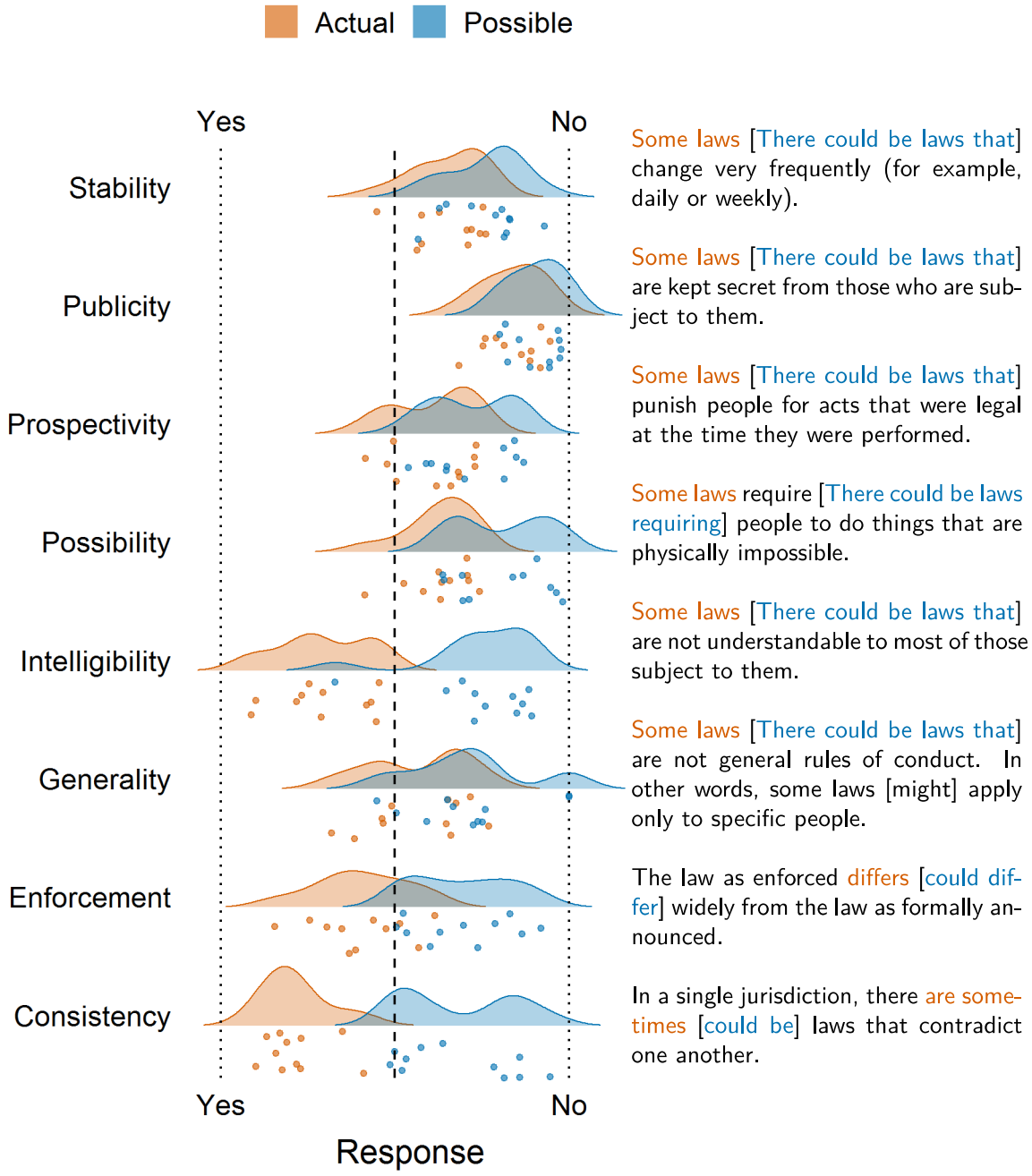


Figure 1. Rain cloud plot: Probability density and scatter plot by procedural principle and condition. Each dot represents a country.

3.2 Exploratory analyses

3.2.1 Manipulation check

Modal language is employed primarily to represent necessity and possibility. For instance, the question "Could there be life on Mercury?" concerns a physical possibility, just as the assertion that "There could not be any married bachelors" expresses a logical

impossibility. Collectively, these are referred to as *alethic* modals, insofar as they purport to assess questions necessity and possibility. In some contexts, however, modal language can also be employed to denote so-called *deontic* properties—i.e., permission and obligation. For instance, the question "Could I sit next to you?" does not, generally speaking, inquire whether the action in question is physically possible (in the alethic sense). Rather, it requests permission from the listener, asking whether the action (i.e., sitting next to the listener) is allowed or forbidden, desirable or undesirable, from the perspective of the listener.

Though these varieties of modality are easily distinguished in thought, natural languages tend to offer rather imperfect ways of doing so—raising the possibility that participants interpreted the statements in our experiment as deontic rather than alethic modals.

To examine this possibility, we probed participants' interpretations of the task through a set of three post-test questions. In mixed-effects linear regression models with country as a random effect, we examined the effect of condition on each interpretation measure (see Table 3). As expected, participants viewed the existential construction ("There *are* laws…") as inviting an empirical assessment about "what laws are *usually like in [their] experience*" (z = -3.96, p = .003). Meanwhile, the modal construction ("There *could be* laws…") was interpreted as describing *both* what laws "must be like in order to count as law" (an *alethic* interpretation; z = 3.38, p = .008) and what laws "should be like ideally, according to [one's] beliefs about right and wrong" (a *deontic* interpretation; z = 3.61, p = .006).

Table 3.

Task interpretation: Marginal means by condition.

	Marginal means [95% CI]		**Fixed effect (Condition)**		**Random effects (Country)**	
	Actual	**Possible**	***z***	***p***	***(int.)***	***(slope)***
Alethic	1.01 [0.73, 1.29]	1.44 [1.23, 1.64]	3.38	.008	0.39	0.37

Deontic	0.71 [0.33, 1.10]	1.24 [1.03, 1.45]	3.61	.006	0.55	0.43
Empirical	1.24 [1.05, 1.43]	0.68 [0.32, 1.04]	-3.96	.003	0.25	0.43

3.2.2 Effects of task interpretation

The measures of participants' interpretation of the task were moderately correlated in each condition (see Table 4). As such, estimating the impact of task interpretation via moderation analyses could be jeopardized by the presence of multicollinearity (but see Supplementary Analysis 3). Instead, to examine whether our primary finding depended on participants' interpretation of the task (and, particularly, of the modal construction in the Possible condition), we conducted *latent profile analyses* (Collins & Lanza, 2009) identifying patterns of interpretation (or latent profiles) across the three items in each condition.

Table 4.

Correlation between interpretation measures in the Actual (below diagonal) and Possible (above diagonal) conditions.

		(1)	**(2)**	**(3)**
(1)	*Alethic*	-	.35 *** [.31, .40]	.48 *** [.44, .52]
(2)	*Deontic*	.48 *** [.44, .52]	-	.15 *** [.09, .20]
(3)	*Empirical*	.03 [-.02, .09]	.01 [-.05, .06]	-

These person-centered analyses revealed that differences in task interpretation were driven by a minority profile in each condition: In the Actual condition, one-in-five (n = 352) participants reported reflecting on the empirical facts about law ($M_{\text{empirical}}$ = 1.59, 95% CI [1.43, 1.74], SD = 1.50), but not its necessary (M_{alethic} = -0.32, 95% CI [-0.53, -0.12], SD = 1.98) or deontic (M_{deontic} = -2.07, 95% CI [-2.16, -1.97], SD = 0.92) properties. Meanwhile,

the majority reported a fairly indiscriminate, hybrid interpretation (1.19 < all *M*s < 1.47, 1.14 < all *SD*s < 1.42). Similarly, a minority in the Possible condition (n = 237) reported focusing on the necessary properties of law (M_{alethic} = 0.56, 95% CI [0.22, 0.67], SD = 2.07; $M_{\text{empirical}}$ = 0.30, 95% CI [-0.06, 0.37], SD = 1.97), but not its deontic properties (M_{deontic} = -1.84, 95% CI [-1.97, -1.71], SD = 1.02)—while again a large majority reported a hybrid interpretation (0.82 < all *M*s < 1.79, 1.05 < all *SD*s < 1.58).

Still, our primary prediction (concerning the effect of condition on endorsement) was borne out even among participants in the minority profiles—who reported a selective interpretation (*Actual:* $\hat{p}$ = .50, 95% $\text{CI}_{\text{asymptotic}}$ [.36, .64]; *Possible:* $\hat{p}$ = .78, 95% $\text{CI}_{\text{asymptotic}}$ [.65, .86]), OR = 3.46, z = 11.65, p < .001. This result rules out one particular explanation for the primary finding: namely, that participants interpreted the modal in a deontic manner, asking themselves whether laws *should* observe the procedural principles in question (or be allowed to violate them); and therefore that the difference between conditions reflects the recognition that laws occasionally violate certain principles that ideally they ought to observe.

4. Discussion

Countries and jurisdictions differ substantially in the extent to which their legal systems observe fundamental principles of the rule of law (World Justice Project, 2020). One might expect that this variation would lead to cultural differences in people's concept of law. Yet, our present findings suggest that there is a striking level of agreement about the procedural constraints of law—observable throughout the adult life span, across highly dissimilar cultures and languages, and in lawyers and laypeople alike. People consistently believe that laws necessarily abide by a series of procedural principles: they could not retroactively punish past conduct, be kept secret, or be incomprehensible to most, for instance. Yet, people also acknowledge that laws in practice violate these very principles.

The legal systems of the eleven countries in the study vary in large and significant ways; the sample includes civil law and common law systems, countries with varying degrees

of religious influence on their laws, and countries with diverse political, representational, and legislative systems. These features are obviously relevant to other important legal and political questions, but it is striking that laypeople—across all of these different legal system—share common intuitions about the concept of law.

4.1 Future directions and limitations

Although our study provided clear evidence of the phenomenon in question, it afforded limited insight into the psychological processes that engender conflicting beliefs about actual versus possible laws.

A growing body of evidence demonstrates mutual influences between descriptive and prescriptive reasoning: Descriptive considerations—such as whether a behavior is common or rare—can sometimes be vested with normative weight, e.g., when people punish non-conformity to group norms (Roberts, Gelman, & Ho, 2017; Roberts, Ho, & Gelman, 2019). Similarly, prescriptive considerations—such as whether a particular conduct is morally good or evil—can shape descriptive judgments, e.g., about whether such conduct is even possible (Phillips & Cushman, 2017; Goldring & Heiphetz, 2020).

Our present findings could be understood as a manifestation of the latter phenomenon: When prompted to reason about whether laws necessarily exhibit certain procedural constraints, most participants also considered whether such constraints would be desirable or undesirable. And yet, in both person-centered and variable-centered analyses (see Supplementary Analysis 3), even those who interpreted the modals in a purely alethic fashion demonstrated a comparable effect. As such, our study provided at least some negative evidence concerning the mechanism: i.e., that the effect is *not* driven by the ambiguity that pervades modality in natural languages.

Since we assessed participants' engagement in prescriptive thinking through their *explicit* reports, this still leaves open unconscious pathways for prescriptive considerations to alethic judgments of necessity and possibility. In previous research, the tendency to interpret

immoral states of affairs as impossible arises mostly strongly in conditions favoring a quick and intuitive assessment (e.g., under time pressure, see Phillips & Cushman, 2017). Therefore, future work could investigate whether a broader, intuitive association between immorality and impossibility undergirds people's endorsement of procedural constraints on law.

A long tradition of research on essentialist thinking has demonstrated that people infer category membership—for instance, whether something is a tiger (Gelman & Wellman, 1991), or a work of art (Liao et al., 2020)—on the basis of abstract, and even unobservable, qualities which constitute its essence (Gelman, 2004; Keil, 1989). The tendency toward *essentialization* has also been observed in the legal context, to explain stigma surrounding criminals (i.e., as criminality is attributed to an intrinsic property of criminals; see Dunlea & Heiphetz, 2020). Our present results can also be seen in this light: i.e., as revealing the qualities that people consider the *essence* of law, and which readily come to mind when considering prototypical category members.

Relatedly, recent work reveals that various abstract concepts, such as love, friendship or art (Del Pinal & Reuter, 2017; Knobe et al., 2013; Leslie, 2015; Liao et al., 2020) also demonstrate certain trademarks of essentialist thinking: Building on the notion of dual character concepts, these studies demonstrate that being a *true* friend or a *true* artist depends on the presence of deeper criteria for category membership, even when the superficial criteria may be absent (Newman & Knobe, 2019; Rose & Nichols, 2020). This body of evidence points toward a promising avenue for further research: Potentially, by comparison to empirical reasoning about actual laws, modal reasoning about possible laws draws attention toward deeper criteria for category membership (i.e., could *true* laws lack these procedural properties?), such as the purpose of laws (Rose & Nichols, 2020) or their moral dimension (Flanagan & Hannikainen, 2020; Knobe et al., 2013).

Finally, we must note that the magnitude of the effect varied substantially across countries, which may imply some degree of cultural variability. Relatedly, our sampling methods differed substantially across sites; so variation in the magnitude of the effect might also reflect differences in sample composition, attentiveness, or even acquiescence (see Heine, Lehman, Peng, & Greenholtz, 2002).

4.2 Conclusion

Cognitive science has made ample progress in drawing the contours of the moral (Cushman et al., 2006; Graham et al., 2009) and economic (Boyer & Pedersen, 2018; Henrich et al., 2005) mind—while comparatively neglecting to investigate the psychological and cultural basis of legal concepts. The present work represents an early step in this research program, and provides evidence of cross-cultural convergence in people's understanding of the nature of law. We reveal a striking degree of universality in beliefs about the quintessential properties of law, despite abundant historical and regional variation in the way actual laws manifest.

References

Awad, E., Dsouza, S., Shariff, A., Rahwan, I., & Bonnefon, J. F. (2020). Universals and variations in moral decisions made in 42 countries by 70,000 participants. *Proceedings of the National Academy of Sciences, 117*(5), 2332-2337.

Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M., Fitzpatrick, S., Gurven, M., ... & Laurence, S. (2016). Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. *Proceedings of the National Academy of Sciences, 113*(17), 4688-4693.

Bates, D., Maechler, M., & Bolker, B. S. Walker (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Boyer, P., & Petersen, M. B. (2018). Folk-economic beliefs: An evolutionary cognitive model. *Behavioral and Brain Sciences, 41*.

Brown, D. E. (1991). *Human universals.* Mc-Graw Hill.

Collins, L. M., & Lanza, S. T. (2009). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences* (Vol. 718). John Wiley & Sons.

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological Science, 17*(12), 1082-1089.

Del Pinal, G., & Reuter, K. (2017). Dual character concepts in social cognition: Commitments and the normative dimension of conceptual representation. *Cognitive Science, 41*, 477-501.

Donelson, R, & Hannikainen, I. R. (2020). The Inner Morality of Law Revisited. *Oxford Studies in Experimental Philosophy: Volume 3, 3*, 6.

Dunlea, J. P., & Heiphetz, L. (2020). Children's and adults' understanding of punishment and the criminal justice system. *Journal of Experimental Social Psychology, 87*, 103913.

Finnis, J. (1980). *Natural law and natural rights*. Oxford University Press.

Flanagan, B., & Hannikainen, I. R. (2020). The Folk Concept of Law: Law Is Intrinsically Moral. *Australasian Journal of Philosophy*, 1-15.

Fuller, L. L. (1964). *The morality of law* (Vol. 21). New Haven: Yale University Press.

Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of personality and social psychology*, *96*(5), 1029.

Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology*, *39*(1), 113-127.

Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002).What's wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect. *Journal of Personality and Social Psychology*, 82, 903–918.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., ... & Henrich, N. S. (2005). "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and brain sciences*, *28*(6), 795-815.

Hoebel, E. A. (1954). *The law of primitive man: a study in comparative legal dynamics*. Cambridge, Mass., Harvard University Press.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development.* MIT Press, Cambridge, MA.

Knobe, J., Prasada, S., & Newman, G. E. (2013). Dual character concepts and the normative dimension of conceptual representation. *Cognition*, *127*(2), 242-257.

Margolis, E., & Laurence, S. (Eds.). (1999). *Concepts: core readings*. Mit Press.

Mikhail, J. (2007). Universal moral grammar: theory, evidence and the future. *Trends in Cognitive Sciences*, 11(4): 143-152.

Mikhail, J. (2009). Is the prohibition of homicide universal? Evidence from comparative criminal law. *Brook. L. Rev.*, *75*, 497.

Nader, L. (1965). The Anthropological Study of Law 1. *American Anthropologist*, *67*(6), 3-32.

Newman, G. E., & Knobe, J. (2019). The essence of essentialism. *Mind & Language*, *34*(5), 585-605.

Ofosu, E. K., Chambers, M. K., Chen, J. M., & Hehman, E. (2019). Same-sex marriage legalization associated with reduced implicit and explicit antigay bias. *Proceedings of the National Academy of Sciences*, *116*(18), 8846-8851.

Phillips, J., & Cushman, F. (2017). Morality constrains the default representation of what is possible. *Proceedings of the National Academy of Sciences*, *114*(18), 4649-4654.

Rose, D., & Nichols, S. (2020). Teleological essentialism: generalized. *Cognitive Science*, *44*(3), e12818.

Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: a mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of personality and social psychology*, *76*(4), 574.

Sznycer, D., & Patrick, C. (2020). The origins of criminal law. *Nature human behaviour*, *4*(5), 506-516.

Thalmayer, A. G., Toscanelli, C., & Arnett, J. J. (2021). The neglected 95% revisited: Is American psychology becoming less American? *American Psychologist*, *76*(1), 116–129.

Author Contributions

IRH and RD conceived the study concept. IRH, KPT, GA, VD, MK, and NS developed and designed the study. All authors were involved in data collection and stimuli development. IRH wrote first draft with substantial contributions from KPT. All authors provided critical revisions and approved the final version of the manuscript for submission.

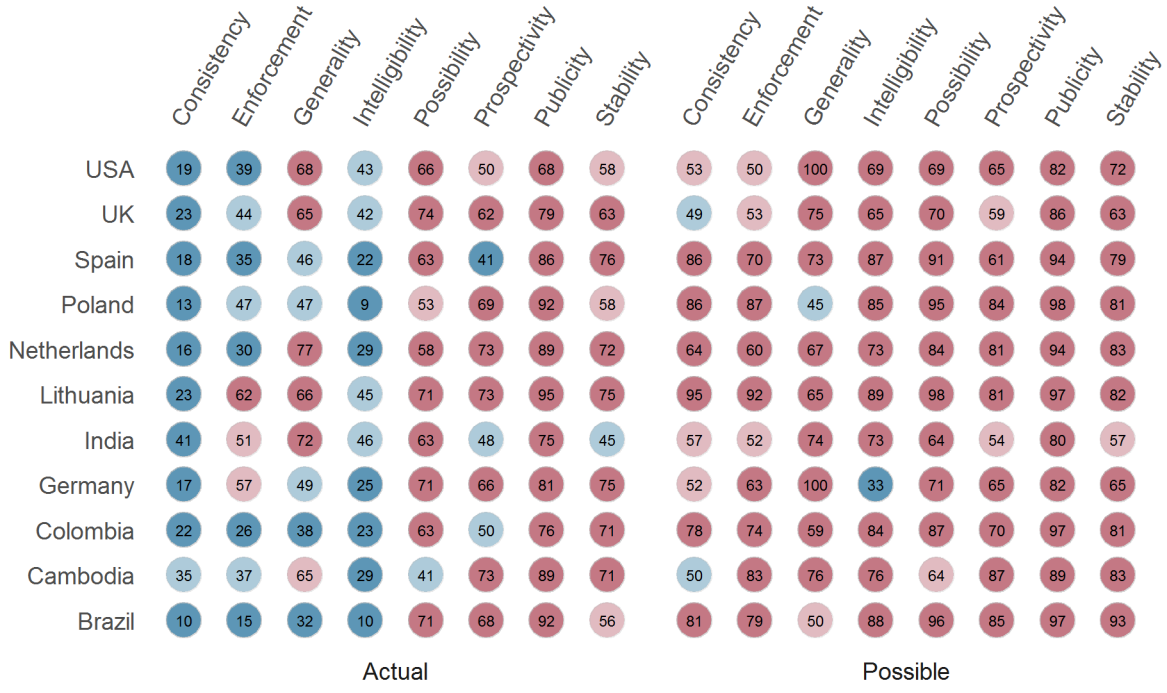
Supplementary Analysis 1: Sub-Group Analyses

We enter a series of demographic measures as covariates in the pre-registered model. We then examine the simple effect of condition at each level of: (1) law background, (2) gender, and (3) age bracket.

	n	Actual	Possible	*OR*	*z*	*p*
Law background						
Law Student	293	.49 [.34, .64]	.82 [.71, .90]	4.93	10.58	< .001
Law Graduate	762	.54 [.39, .68]	.77 [.65, .86]	2.81	11.34	< .001
Neither	1929	.55 [.41, .69]	.83 [.73, .90]	3.96	23.55	< .001
Gender						
Female	1480	.54 [.39, .67]	.81 [.70, .88]	3.85	19.64	< .001
Male	1491	.55 [.40, .68]	.82 [.71, .89]	3.72	19.76	< .001
Non-binary	16	.44 [.21, .70]	.83 [.63, .93]	6.45	2.85	.004
Age bracket						
24 or younger	682	.47 [.33, .62]	.82 [.71, .89]	4.98	17.09	< .001
25 to 34	794	.52 [.38, .67]	.78 [.67, .87]	3.27	13.78	< .001
35 to 44	571	.59 [.43, .72]	.82 [.71, .89]	3.13	11.20	< .001
45 to 54	369	.60 [.45, .73]	.84 [.73, .90]	3.40	9.54	< .001
55 to 64	192	.65 [.50, .78]	.87 [.78, .93]	3.57	7.02	< .001
65 or older	127	.65 [.49, .78]	.90 [.81, .94]	4.69	6.74	< .001

Supplementary Analysis 2. Absolute Endorsement by Condition

In the Results section, we documented a *comparative* effect, such that participants were more likely to report that laws would have to observe various procedural constraints than that they actually do. But did participants view these procedural principles as necessary in an *absolute* sense? As shown in Supplementary Figure 1, looking separately by country, participants in the Actual condition were unsure whether laws observed each of the principles (48 out of 88, or 55%), whereas participants in the Possible condition tended to agree that laws could not violate any of the procedural principles (84 out of 88, or 95%)—potentially indicating that people's beliefs about possible laws are shaped to a weaker extent by their first-hand experience with real laws.



Supplementary Figure 1. Absolute endorsement (%) by country and principle. Endorsement is highlighted in red, and rejection in blue. Dark shades represent statistical significance in a proportion test against the null of a uniform distribution.

Supplementary Analysis 3: Effects of Interpretation via Multiple Regression

Entering the interpretation measures in our primary regression model revealed both main effects of interpretation, as well as interactions with condition. However, the variance inflation factors for the predictors in the model ranged between 2.07 and 2.97—raising some concerns about multicollinearity, and suggesting that the corresponding coefficients should be interpreted with some caution. Alethic interpretations predicted endorsement in both conditions (*Actual*: OR = 1.16, 95% CI [1.11, 1.21], z = 6.72; *Possible*: OR = 1.15, 95% CI [1.10, 1.21], z = 5.99; ps < .001). Meanwhile, empirical interpretations were tied to reduced endorsement of the principles—though only in the Possible condition (*Possible*: OR = 0.93, 95% CI [0.90, 0.97], z = -3.17, p = .002; *Actual*: OR = 1.03, 95% CI [0.98, 1.07], z = 1.23, p = .22). Lastly, as might be expected, deontic interpretations increased endorsement of the principles in the Possible condition only (OR = 1.13, 95% CI [1.09, 1.18], z = 6.07, p < .001; *Actual*: OR = 1.01, 95% CI [0.97, 1.05], z = 0.55, p = .58). Critically, the effect of condition remained highly significant in the aforementioned model (see Supplementary Table 2). At a broad level, these analyses reveal that the effect of condition was remarkably robust to differences in task interpretation—though indeed individual interpretations predicted responses, especially in the presence of the modal auxiliary (i.e., in the Possible condition).

Supplementary Table 2.

Effects of interpretation on endorsement: Mixed-effects, logistic regression.

	Additive Model 1 (*AIC* = 26346)			**Interaction Model 2 (*AIC* = 26847)**			
Fixed effects	**OR**	***z***	***p***	**OR**	***z***	***p***	**VIF**
Condition	3.11 [1.83, 5.29]	4.20	< .001	3.38 [2.98, 3.84]	18.72	< .001	2.07
Empirical	0.96 [0.93, 0.99]	-2.46	.014	0.96 [0.93, 0.99]	-2.74	.006	2.07

Alethic	1.16 [1.12, 1.20]	9.00	< .001	1.16 [1.13, 1.19]	9.41	< .001	2.18
Deontic	1.07 [1.04, 1.10]	4.66	< .001	1.08 [1.05, 1.11]	5.32	< .001	2.22
Empirical×Condition	-	-	-	0.86 [0.81, 0.91]	-5.01	< .001	2.33
Alethic×Condition	-	-	-	1.01 [0.95, 1.08]	0.35	.73	2.97
Deontic×Condition	-	-	-	1.11 [1.05, 1.17]	3.69	< .001	2.55
intercept	1.78 [1.00, 3.18]	1.86	.050	1.76 [0.98, 3.15]	1.90	.057	-

Random effects		**SD**	**# Groups**	**ICC**
Participant:Country	Intercept	0.89	3023	.165
Country	Intercept	0.28	11	.015
Principle	Intercept	0.81	8	.226

Note. Condition is effect-coded (-0.5 = *Actual* and 0.5 = *Possible*) so the empirical, alethic, and deontic parameters represent main effects in both models (when Condition = 0).