

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

How to make a reasonable person

Permalink

<https://escholarship.org/uc/item/3dg945t9>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors

Wu, Sarah A

Zhang, Siying

Gerstenberg, Tobias

Publication Date

2026

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

How to make a reasonable person

Sarah A. Wu^{1*}, Siying Zhang² & Tobias Gerstenberg¹

¹Department of Psychology, Stanford University

²Department of Psychology, University of Washington

*sarahawu@stanford.edu

Abstract

Many legal decisions rely on evaluating how a “reasonable person” would have acted, but what defines this standard? Here, we offer an experimental jurisprudence perspective by investigating what subjective features come to mind when reasoning about a reasonable person. In Experiment 1, we examined how people conceptually organize demographic, dispositional, and action features along dimensions of relevance, controllability, and normality. In Experiment 2, we used these dimensions to predict what features shape judgments about reasonableness and outcomes. We found that participants “undid” harmful outcomes by focusing on relevant, abnormal actions; that they constructed a reasonable person from the defendant by preserving normal attributes while changing abnormal, controllable ones; and that they endorsed subjective standards based on features that were relevant and normal, but relatively uncontrollable, consistent with theories of blame. Together, these findings map a template of a reasonable person and provide a descriptive foundation to inform legal theory.

Keywords: reasonable person; experimental jurisprudence; legal reasoning; psychology and law

Introduction

A great deal of American law rests on the notion of a “reasonable person”. To determine whether someone negligently caused an injury (Zipursky, 2015), harassed someone else (Abrams, 1998; Schultz, 1998), or even committed murder in purported self-defense (Lee, 2003), jurors are often instructed to consider how a reasonable person would have acted in the defendant’s situation. Legal scholars have long debated whether reasonableness should be judged according to a uniform external benchmark (an objective standard) or based on the defendant’s particular features and capacities (a subjective standard) (Ben-Shahar & Porat, 2021; Fletcher, 1971; Heller, 1998; Lee, 2003; McCormick, 1999; Moran, 2003, 2010; Nourse, 2008; Scalet, 2003; Westen, 2008). On the one hand, a subjective standard allows jurors to recognize individual circumstances and account for diversity in lived experience. On the other hand, an objective standard encourages predictability and neutrality, and can help safeguard against any personal prejudices (Lee, 2003).

In practice, many courts employ a hybrid standard of reasonableness that allows for the subjectification of certain features, such as physical disabilities and professional experience, but not necessarily others, such as age (Northrop and Rozan, 2017; e.g. *Dellwo v. Pearson*, 1961), gender (Gutek and O’Connor,

1995; Perry et al., 2004; Shoenfelt et al., 2002; e.g. *Ellison v. Brady*, 1991), cultural background (Renteln, 2005; e.g. *People v. Wu*, 1991) or psychological conditions (Faigman, 1986; Moran, 2010; e.g. *People v. White*, 1981; *State v. Kelly*, 1984). When and why should some features matter, but not others? In this paper, we offer an experimental jurisprudence perspective on this debate (Sommers, 2021; Tobia, 2022). We investigate the *ordinary* concept of the reasonable person – how ordinary people, who act as jurors, intuitively construct such a person in their minds.

Ordinary reasonableness

Prior work on how people judge reasonableness suggests that they understand it to be more of an empirical concept (i.e. what people tend to do) rather than an economic one (i.e. what is welfare-maximizing) (Jaeger, 2021; Martinez & Tobia, 2023). Furthermore, this empirical concept seems to be a hybrid of what is statistically average and what is ideal (Tobia, 2018; Tobia et al., 2025; but see Baumgartner and Kneer, 2025). Judgments of reasonableness are influenced by many factors, including knowledge of the outcome (Kneer, 2022) and inferences about the mental states of the agents involved (Margoni & Brown, 2023). There are also considerable individual differences: what people think is reasonable is shaped by their unique identities and social contexts (Spruill & Lewis, 2023), and often driven by standards of their own behavior (Alicke & Weigel, 2021; Simpson et al., 2020). While these studies offer insight into what concepts are at play when people evaluate reasonableness, they do not address what specific features matter when imagining a reasonable person in a particular defendant’s situation. Empirical evidence on this question could help inform legal debates about which features courts should permit in jury instructions, which have thus far relied largely on normative arguments.

The present study

Here, we explore what features come to people’s minds when reasoning about a hypothetical reasonable person. We designed fictional legal cases involving characters whose actions resulted in harmful outcomes (Figure 1A). The characters had diverse attributes, and the cases covered different domains where the reasonable person standard is commonly applied, such as negligence, self-defense, and provocation. For each case, we specified features about the defendant including demographic characteristics (e.g. gender), dispositions (e.g.

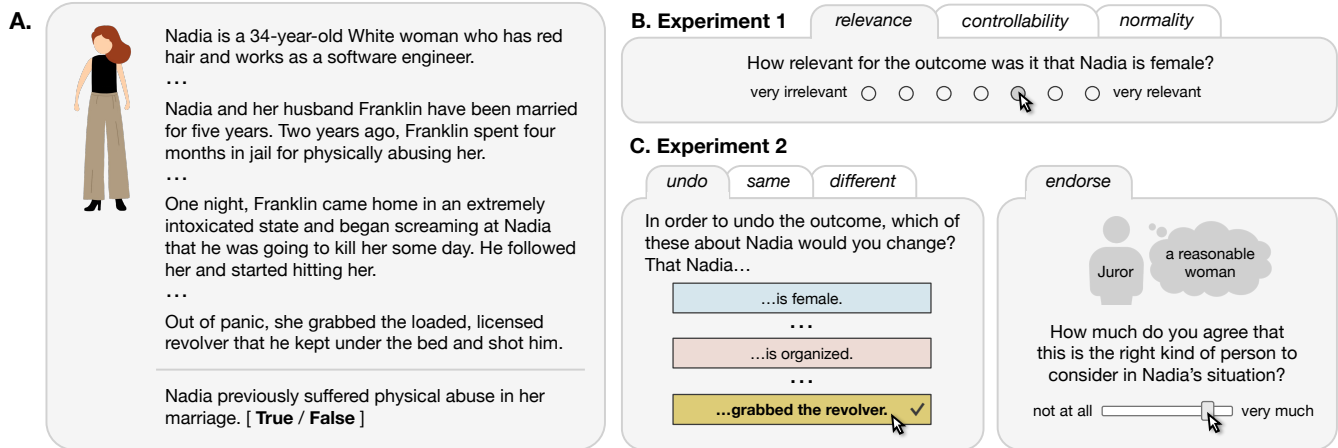


Figure 1: **Overview of experiments.** (A) Example case involving domestic abuse and self-defense. (B) In Experiment 1, participants rated the relevance, controllability, or normality of features about each defendant on a Likert scale. (C) In Experiment 2, the conditions had one of two designs. Participants either selected features they would change to undo the outcome or selected features in response to prompts about what would be the same as or different from a reasonable person in the defendant's situation (left), or made judgments about hypothetical subjective reasonable person standards (right). The features were presented in random orders, and the coloring of the different feature types was not shown in the experiment (just for illustrative purposes here).

preferences or personality traits), and actions the defendant took (e.g. "drove over the speed limit").

In Experiment 1, we examined how these features are judged on three dimensions central to moral and legal judgments (Figure 1B). The first dimension was how *relevant* the feature was to the outcome, reflecting the law's emphasis on proximate cause and the importance of causation in people's moral reasoning (e.g. Hart & Honoré, 1959; Knobe & Shapiro, 2021; Lagnado & Gerstenberg, 2017; Waldmann et al., 2012). The second dimension was how much *control* or agency the defendant had over the feature, as perceived control tends to increase judgments of blame (e.g. Alicke, 2000; Gilbert et al., 2015; Giroto et al., 1991; Weiner, 1995). The last dimension was how *normal* the feature was, since normality is strongly tied to causal and moral judgments (e.g. Hitchcock & Knobe, 2009; Kahneman & Miller, 1986; Kominsky & Phillips, 2019; Kominsky et al., 2015). Because these three factors are well-established in prior literature, we used them as a theoretically motivated starting point to investigate whether they might play an analogous role in shaping intuitions about reasonableness. In a between-subjects design, participants rated each feature for each case on one of these dimensions.

In Experiment 2, we used these dimensions to predict which features play a role in various judgments about the cases (Figure 1C). In a between-subjects design, participants made one of four judgments. In the *same* and *different* conditions, participants identified which features would be the same or would differ between the defendant and a reasonable person in the defendant's situation. This allowed us to probe two possible strategies for mentally constructing the reasonable person: starting with the defendant and removing any unreasonable

features, or starting from "scratch" and adding reasonable features. In the *endorse* condition, participants evaluated whether specific demographic and dispositional features of the defendant should define the reasonable person standard in their situation. Finally, in the *undo* condition, participants indicated which features they would change to prevent the outcome. This allowed us to distinguish between what makes someone (un)reasonable and what merely caused harm.

Experiment 1: Feature dimensions

In this experiment, we gathered participants' judgments about how relevant, controllable, and normal the features of each defendant were.

Methods

All experiments were pre-registered on the Open Science Framework and programmed using jsPsych (de Leeuw et al., 2023).¹

Participants We recruited 155 participants from Prolific who were all native English speakers residing in the US, and compensated each at a rate of \$12/hour. 5 participants were excluded for failing attention checks, leaving a final sample size of 150 (age: $M = 40$, $SD = 14$; gender: 70 female, 75 male, 2 non-binary, 3 undisclosed; race: 108 White, 19 Black/African American, 8 Asian, 1 American Indian/Alaska Native, and 14 undisclosed). Participants were assigned to either the *relevance*, *controllability*, or *normality* condition with $N = 50$ in each.

¹All data and materials are available at: https://github.com/cicl-stanford/reasonable_person_cogsci. We acknowledge the help of AI tools in programming the experiments.

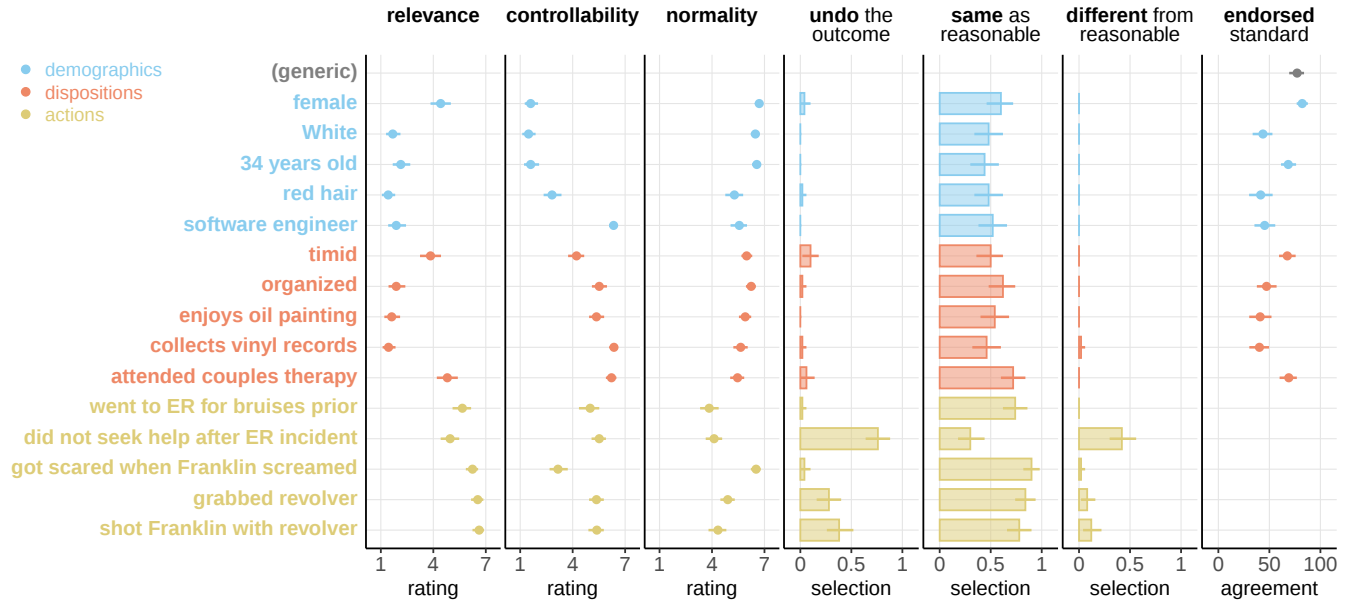


Figure 2: **Results for the example case from Figure 1.** The y-axis lists all the features of the defendant. The x-axes show mean Likert ratings of relevance, controllability, and normality; proportion selected to undo the outcome, and to be the same or different in a reasonable person; and mean endorsement ratings (for demographics and dispositions only, as well as a generic reasonable person). The first three measures are from Experiment 1, and the last four are from Experiment 2. Note that all conditions were between-subjects. Error bars are 95% confidence intervals.

Procedure Participants read 11 fictional scenarios that were each presented across three pages. The first page displayed the story, the second page hid the story and asked a single true/false comprehension question, and the third page displayed the story again along with the test questions (see Figure 1 for an example). For each story, participants were asked to rate features about the defendant on a 7-point Likert scale. Depending on the condition, they were asked “How relevant for the outcome was it that [name] [feature]?” with the scale ranging from “very irrelevant” to “very relevant”, “How much control did [name] have over [feature]?” ranging from “none”

to “very much”, or “How normal was it that [name] [feature]?” ranging from “very abnormal” to “very normal”. The experiment took an average of 24 minutes (SD = 11) to complete.

Design Scenarios were adapted from prior studies and actual legal cases (Jaeger, 2021; Kneer, 2022; Simpson et al., 2020). For each scenario, we designed 15 features about the defendant consisting of five demographics, five dispositions, and five actions the defendant took. The order of the scenarios was randomized for each participant. Within each scenario, the order of the features was randomized except that the actions were kept together in chronological order. Participants were excluded if they failed more than two comprehension questions out of the 11 total.

Results and discussion

Ratings for the features of an example case are shown in the left three columns of Figure 2. This case described a woman in an abusive marriage who fatally shot her husband when he began assaulting her. Participants’ judgments spanned the entire scale for all three dimensions, and there was noticeable variation even within different feature types. For example, the defendant being a woman was judged as much more relevant than any of the other demographics. Meanwhile, her getting scared when her husband screamed was judged to be relatively less controllable and more normal than the other actions.

Overall, across all stories, there were systematic trends in how relevant, controllable, and normal participants rated the different feature types. To account for the non-independence

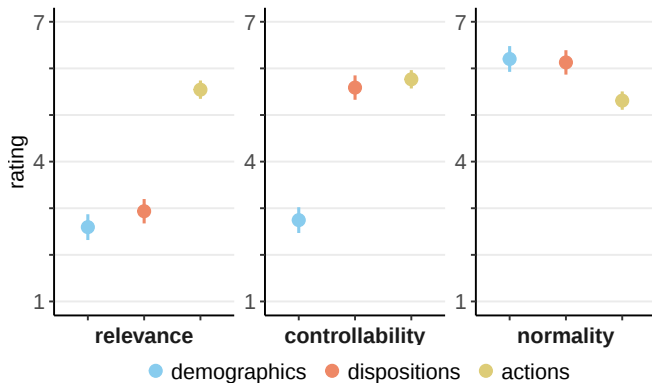


Figure 3: **Experiment 1 results.** Mean ratings for the different feature types across all scenarios. Error bars show 95% confidence intervals.

of observations within participants, we used a linear mixed effects model with a random intercept and random slopes (by participant) to estimate marginal means and confidence intervals. First, on relevance, participants rated actions as significantly more relevant to the outcome ($M = 5.54$, 95% CI [5.35, 5.74]) than demographic characteristics ($M = 2.59$, 95% CI [2.32, 2.87], $z = 20.7$, $p < 0.001$) and dispositions ($M = 2.94$, 95% CI [2.67, 3.20], $z = 22.7$, $p < 0.001$). Second, on controllability, participants rated the controllability of actions ($M = 5.77$, 95% CI [5.57, 5.96]) and dispositions ($M = 5.59$, 95% CI [5.33, 5.85]) as significantly higher than that of demographics ($M = 2.75$, 95% CI [2.47, 3.02]; vs. actions: $z = 21.2$, $p < 0.001$; vs. dispositions: $z = 28.2$, $p < 0.001$). Finally, on normality, participants rated actions ($M = 5.31$, 95% CI [5.11, 5.51]) as significantly less normal than demographics ($M = 6.20$, 95% CI [5.93, 6.48], $z = -6.3$, $p < 0.001$) and dispositions ($M = 6.13$, 95% CI [5.87, 6.39], $z = -7.1$, $p < 0.001$), though mean judgments for all three feature types were above the midpoint. Together, these results suggest that participants viewed the outcomes of the scenarios as largely unrelated to most ordinary demographic and dispositional traits of the defendants, seeing them instead as hinging primarily on how defendants acted in unexpected ways. While this overall pattern is unsurprising, there was also meaningful variation within feature types and across different stories.

Experiment 2: Reasonableness judgments

The results of Experiment 1 highlight systematic patterns in how people conceptually organize various features about a person, specifically along the dimensions of relevance, controllability, and normality. Next, we investigated how these dimensions predict which features matter most in judgments about outcomes and reasonableness.

Methods

Participants We recruited 211 participants from Prolific who were all native English speakers residing in the US, and compensated each at a rate of \$12/hour. 11 participants were excluded for failing attention checks, leaving a final sample size of 200 (*age*: $M = 40$, $SD = 14$; *gender*: 103 female, 91 male, 6 non-binary; *race*: 152 White, 15 Black/African American, 17 Asian, 2 American Indian/Alaska Native, 3 multiracial, and 11 undisclosed). Participants were assigned to the *undo*, *same*, *different*, or *endorse* condition with $N = 50$ in each.

Procedure The procedure was similar to that of Experiment 1. In the *undo* condition, participants were asked to select any number of features in each scenario in response to the question “In order to undo the outcome, which of these about [name] would you change? That [name] . . .” (Figure 1C). All features were initially unselected. In the *same* condition, the question read “Imagine a reasonable person in the same situation as [name]. A reasonable person would also have . . .”. In the *different* condition, the question read

“Imagine a reasonable person in the same situation as [name]. A reasonable person would not have . . .”.

The *endorse* condition had a different design (Figure 1C). Participants were presented with jurors who imagined different hypothetical reasonable person standards, and were asked “How much do you agree that this is the right kind of person to consider in [name]’s situation?”. They responded using continuous sliders with endpoints labeled “not at all” (0) and “very much” (100). Overall, the experiment took an average of 19 minutes ($SD = 9$) to complete.

Design The study design was identical to that of Experiment 1, except that in the *endorse* condition, we excluded actions and only presented hypothetical standards based on demographics or dispositions. We also asked participants in this condition to rate their agreement with a “reasonable generic person” standard, which was always presented first before a randomized ordering of the rest of the features.

Results

The results for an example case are shown in the right four columns of Figure 2. Participants generally selected more features to be the same in a reasonable person, compared to what might be different or changed to undo the outcome. They primarily selected actions to undo the outcome or be different, the main one being seeking help after an ER incident in both conditions. Across the demographic and disposition features, participants endorsed their inclusion in a subjective reasonable person standard to varying degrees (≈ 40 –80%).

Figure 4 shows aggregate findings. Figure 4A illustrates the distribution of features selected by participants in the *undo*, *same*, and *different* conditions in the first three columns. The last column (top-right panel) shows the distribution of endorsement judgments for the generic standard and for standards involving demographics or dispositions. Figure 4B shows how well relevance, controllability, and normality predict participants’ selections and agreement ratings. We estimated these coefficients by fitting Bayesian mixed effects models to participants’ responses using mean ratings from Experiment 1 as fixed effects. We fit a linear regression for the *endorse* condition, and logistic regressions for the other three. All models included a random intercept and random slopes by predictor for each participant. Figure 4C compares participants’ responses with regression model predictions.

We describe the results of each condition in turn.

What features do people change to *undo* the outcome?

Participants did not choose many features to undo the outcome, but of the ones they did, the vast majority were actions (Figure 4A). We found a positive effect of relevance and a negative effect of normality on the features selected. Specifically, each one-point increase in relevance rating more than doubled the odds of the feature being selected (odds ratio (OR) = 2.51, 95% credible interval (CrI) [2.17, 2.92]), and each one-point increase in normality rating more than halved those odds (OR = 0.41, 95% CrI [0.36, 0.47]). The effect of controllability was

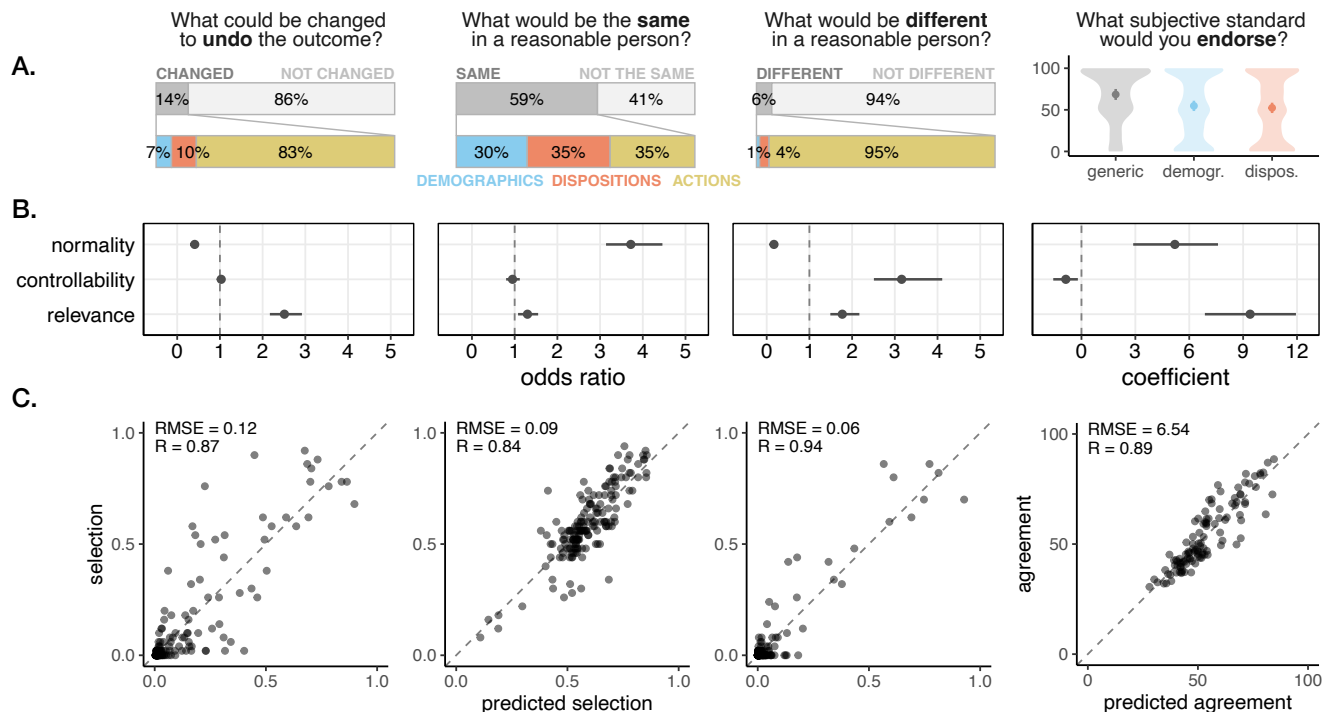


Figure 4: **Experiment 2 results.** Each column represents a set of results from one condition: *undo*, *same*, *different*, and *endorse*. (A) Proportion of features selected, along with breakdown of selections across feature types, in the first three conditions. Distribution of judgments for different types of standards (generic person, demographics, dispositions) in the *endorse* condition. (B) Estimates of odds ratios and coefficients, along with 95% credible intervals, of the effects of normality, controllability, and relevance on the features selected or endorsed. (C) Scatterplots showing mean selection and agreement rates against regression model predictions.

not credible. A regression model based on these dimensions predicted participants' choices well (Figure 4C).

How would the reasonable person be *same* or *different*?

When imagining a defendant compared to a reasonable person in their situation, participants noted many features as the same (59%) and changed very few to be different (6%). The features that were carried over were fairly evenly distributed among demographics, dispositions, and actions, while those that were changed were overwhelmingly actions (95%). Selections in the *same* and *different* conditions were moderately negatively correlated ($r = -0.60$). Between the *same* and *undo* conditions, they were also negatively correlated ($r = -0.40$), but between the *different* and *undo* conditions they were strongly positively correlated ($r = 0.84$).

We found that normality was the strongest predictor of how likely a feature was to be carried over from the defendant to a reasonable person (OR = 3.72, 95% CrI [3.14, 4.46]). Relevance had a small positive effect (OR = 1.30, 95% CrI [1.08, 1.55]). Meanwhile, features were more likely to be changed between the defendant and a reasonable person when they were more relevant (OR = 1.77, 95% CrI [1.49, 2.17]) and more controllable (OR = 3.16, 95% CrI [2.51, 4.11]), and less likely to be changed when they were more normal

(OR = 0.17, 95% CrI [0.13, 0.22]). When entered into regression models, these three dimensions strongly predicted participants' overall choices (Figure 4C).

What features of a reasonable person do people *endorse*?

Participants generally endorsed a generic reasonable person standard across all scenarios ($M = 68$, 95% CI [66, 71]). There was more variability in participants' judgments for subjective standards involving demographics ($M = 55$, 95% CI [53, 56]) and dispositions ($M = 52$, 95% CI [51, 54]). Both distributions were bimodal, with many responses clustered at the extreme ends of the scale (Figure 4A). We found credible positive effects of normality ($\beta = 5.20$, 95% CrI [2.88, 7.60]) and relevance ($\beta = 9.39$, 95% CrI [6.87, 11.94]), and a small credible negative effect of controllability ($\beta = -0.87$, 95% CrI [-1.57, -0.20]) on participants' responses. A regression model using these three dimensions predicted mean judgments well (Figure 4C).

Discussion

The results of Experiment 2 reveal consistent, differentiated patterns across judgments about reasonableness and outcomes. In the *undo* condition, participants tried to reverse harmful outcomes by focusing on relevant and atypical ac-

tions taken by the defendant. This aligns with prior research showing that people tend to imagine counterfactual thoughts to abnormal events over normal ones (Kahneman & Miller, 1986; Kahneman & Tversky, 1982; Macrae et al., 1993; Petrocelli et al., 2011). The actions participants chose to undo the outcome were also strongly correlated with those that differed between the defendant and a reasonable person. In Figure 2, we see that the top selections in the *undo* and *different* conditions are roughly the same. Intuitively, people imagine altering the defendant in the ways that led them to bring about (unreasonable) harm in the first place. A key difference is that controllability was a credible predictor for features to change, but not features to undo. One possible explanation is that participants were reluctant to evaluate defendants on the basis of factors outside of their control, yet they still acknowledged those factors as pivotal to the outcome.

When constructing a reasonable person from the defendant, participants used the defendant's ordinary background characteristics as a baseline and selectively modified their abnormal features. Interestingly, many features were designated as neither the same nor different, suggesting some form of non-commitment (Bigelow et al., 2023). That is, people may not be mentally simulating a reasonable person as a complete individual, but rather a minimal representation that includes only the pieces necessary for the specific evaluation at hand (Ho et al., 2022; Johnson-Laird, 2010). Legal scholars have touched on this pluralistic notion of the reasonable person, noting that the standard reflects not a specific identity, but more so a "vehicle" through which different features can be taken into consideration depending on context (Fletcher, 1985; Gardner, 2019; Jeutner, 2024; Moran, 2003; Stern, 2024).

One of the more striking findings here is the polarity in participants' endorsements of various subjective reasonable person standards (Figure 4A). While participants tended to agree on applying a generic person standard across scenarios, they were deeply split on whether to include demographic and dispositional features. Future analysis could determine whether this bimodality stems from variation between features, or between participants *within* features. Such individual differences can arise from participants' unique identities and experiences (Spruill & Lewis, 2023). Future work could also investigate how well participants' judgments can be predicted by their own demographic traits.

General discussion

The reasonable person standard is a cornerstone of American law, yet it remains the subject of much legal debate (e.g. Moran, 2010; Nourse, 2008; Scalet, 2003). Across two experiments, we provide empirical insight into how people mentally construct the demographic, dispositional, and action features of a reasonable person. In Experiment 1, we found that participants systematically organized these features along the dimensions of relevance, controllability, and normality. Actions were judged to be highly relevant and controllable across scenarios, while demographics and dispositions fell more under

background norms, although there were meaningful exceptions across cases. In Experiment 2, we demonstrated how these dimensions guide the features that participants chose to "undo" harmful outcomes, construct a reasonable person from a defendant's traits, and endorse as a subjective standard.

Our results align with prior theories of blame, which similarly emphasize the role of causality, controllability, and normality in moral judgment (Alicke, 2000; Malle et al., 2014; Shaver, 1985; Weiner, 1995). Judgments of blame and reasonableness may draw on shared cognitive processes, such as the tendency to focus counterfactual thinking on abnormal, controllable aspects of an event (Gavanski & Wells, 1989; Kahneman & Miller, 1986). This connection is strongest in the *undo* condition, where participants focused on causally relevant, abnormal features to change the outcome. Meanwhile, the *endorse* condition maps most closely to blame attribution, where both normality and relevance are important in how we hold others responsible for their actions.

Experimental jurisprudence This work contributes to the growing movement of experimental jurisprudence (Sommers, 2021; Tobia, 2022), which seeks to ground legal theory in empirical evidence of how people actually understand legal concepts such as reasonableness (Jaeger, 2023; Tobia et al., 2025; Votruba, 2013). Our findings inform ongoing debates about objective and subjective standards, suggesting that the hybrid reasonable person standard often used in courts reflects basic psychological intuitions about what is fair to hold people responsible for. However, critically, there may still be misalignments between the specific subjective features that are legally permitted in jury instructions and those that people intuitively believe should be included. Policymakers could use data of the kind collected here to audit jury instructions for psychological validity. The bimodality in endorsement judgments also suggests that individual differences in standards of reasonableness may be a source of juror disagreement that future legal theory could address.

Limitations and future directions Our experiments relied on a relatively small set of handcrafted vignettes and features. Future research should explore a broader array of scenarios to test the generalizability of these effects. Moreover, while the dimensions of relevance, controllability, and normality together account well for participants' judgments (Figure 4C), other factors may matter too. For example, the severity of the harm may play a role in shifting what features participants consider reasonable or inflating judgments about relevance and abnormality (Kneer, 2022).

Finally, as AI systems become increasingly sophisticated and integrated into high-stakes settings, some researchers have raised the possibility of using them to assist or even automate legal decisions (Arbel, 2025; Ben-Shahar & Porat, 2021). If AI systems are to determine standards of reasonableness, it is critical to ensure that they are aligned with human intuitions. Our work provides a descriptive foundation that can inform the development of such systems alongside legal theory.

Acknowledgments

We thank Kevin Tobia and members of the Causality in Cognition Lab for helpful discussion and feedback. SAW was supported by a Stanford Interdisciplinary Graduate Fellowship. TG was supported by grants from the Stanford Institute for Human-Centered Artificial Intelligence (HAI) and from the Cooperative AI Foundation.

References

- Abrams, K. (1998). New jurisprudence of sexual harassment. *Cornell Law Review*, 83(5), 1169–1230.
- Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin*, 126(4), 556–574.
- Alicke, M. D., & Weigel, S. H. (2021). The reasonable person standard: Psychological and legal perspectives. *Annual Review of Law and Social Science*, 17, 123–138.
- Arbel, Y. A. (2025). The silicon reasonable person: Can AI predict how ordinary people judge reasonableness?
- Baumgartner, L., & Kneer, M. (2025). The meaning of 'reasonable': Evidence from a corpus-linguistic study. In K. Tobia (Ed.), *The Cambridge Handbook of Experimental Jurisprudence*. Cambridge University Press.
- Ben-Shahar, O., & Porat, A. (2021). *Personalized law: Different rules for different people*. Oxford University Press.
- Bigelow, E. J., McCoy, J. P., & Ullman, T. D. (2023). Non-commitment in mental imagery. *Cognition*, 238, 105498.
- de Leeuw, J. R., Gilbert, R. A., & Luchterhandt, B. (2023). jsPsych: Enabling an open-source collaborative ecosystem of behavioral experiments. *Journal of Open Source Software*, 8(85), 5351.
- Dellwo v. Pearson, 259 Minn. 452 (1961).
- Ellison v. Brady, 924 F. 2d 872 (9th Cir. 1991).
- Faigman, D. L. (1986). The battered woman syndrome and self-defense: A legal and empirical dissent. *Virginia Law Review*, 72(3), 619–647.
- Fletcher, G. P. (1971). Theory of criminal negligence: A comparative analysis. *University of Pennsylvania Law Review*, 119(3).
- Fletcher, G. P. (1985). The right and the reasonable. *Harv. L. Rev.*, 98, 949.
- Gardner, J. (2019). The many faces of the reasonable person. In *Torts and Other Wrongs* (1st ed., pp. 271–303). Oxford University Press/Oxford.
- Gavanski, I., & Wells, G. L. (1989). Counterfactual processing of normal and exceptional events. *Journal of Experimental Social Psychology*, 25(4), 314–325.
- Gilbert, E. A., Tenney, E. R., Holland, C., & Spellman, B. A. (2015). Counterfactuals, control, and causation: Why knowledgeable people get blamed more. *Personality and Social Psychology Bulletin*, 41(5), 643–658.
- Giroto, V., Legrenzi, P., & Rizzo, A. (1991). Event controllability in counterfactual thinking. *Acta Psychologica*, 78(1), 111–133.
- Gutk, B. A., & O'Connor, M. (1995). The empirical basis for the reasonable woman standard. *Journal of Social Issues*, 51(1), 151–166.
- Hart, H. L. A., & Honoré, T. (1959). *Causation in the law*. Clarendon Press.
- Heller, K. J. (1998). Beyond the reasonable man? A sympathetic but critical assessment of the use of subjective standards of reasonableness in self-defense and provocation cases. *American Journal of Criminal Law*, 26(1).
- Hitchcock, C., & Knobe, J. (2009). Cause and Norm. *Journal of Philosophy*, 106(11), 587–612.
- Ho, M. K., Abel, D., Correa, C. G., Littman, M. L., Cohen, J. D., & Griffiths, T. L. (2022). People construct simplified mental representations to plan. *Nature*, 606(7912), 129–136.
- Jaeger, C. B. (2021). The empirical reasonable person. *Alabama Law Review*, 72(4), 71.
- Jaeger, C. B. (2023). Reasonableness from an experimental jurisprudence perspective. In *The Cambridge Handbook of Experimental Jurisprudence*.
- Jeutner, V. (2024). *The reasonable person: A legal biography*. Cambridge University Press.
- Johnson-Laird, P. N. (2010). Mental models and human reasoning. *Proceedings of the National Academy of Sciences*, 107(43), 18243–18250.
- Kahneman, D., & Miller, D. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93(2), 136–153.
- Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In *Judgment Under Uncertainty* (1st ed., pp. 201–208). Cambridge University Press.
- Kneer, M. (2022). Reasonableness on the Clapham Omnibus: Exploring the outcome-sensitive folk concept of reasonable. In *Judicial Decision-Making* (pp. 25–48, Vol. 14). Springer International Publishing.
- Knobe, J., & Shapiro, S. (2021). Proximate cause explained: An essay in experimental jurisprudence. *The University of Chicago Law Review*, 88(1), 165–236.
- Kominsky, J. F., & Phillips, J. (2019). Immoral professors and malfunctioning tools: Counterfactual relevance accounts explain the effect of norm violations on causal selection. *Cognitive Science*, 43(11).
- Kominsky, J. F., Phillips, J., Gerstenberg, T., Lagnado, D., & Knobe, J. (2015). Causal superseding. *Cognition*, 137, 196–209.
- Lagnado, D. A., & Gerstenberg, T. (2017). Causation in legal and moral reasoning. In *The Oxford Handbook of Causal Reasoning* (pp. 565–602, Vol. 1). Oxford University Press.
- Lee, C. (2003). *Murder and the reasonable man: Passion and fear in the criminal courtroom*. New York University Press.
- MacCormick, N. (1999). Reasonableness and objectivity. *Notre Dame Law Review*, 74(5), 1575.
- Macrae, C. N., Milne, A. B., & Griffiths, R. J. (1993). Counterfactual thinking and the perception of criminal behaviour. *British Journal of Psychology*, 84(2), 221–226.

- Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *The Journal of Psychology*, 25(1), 147–186.
- Margoni, F., & Brown, T. R. (2023). Jurors use mental state information to assess breach in negligence cases. *Cognition*, 236, 105442.
- Martinez, E., & Tobia, K. (2023). What do law professors believe about law and the legal academy? *Georgia Law Journal*, 112(1).
- Moran, M. (2003). *Rethinking the reasonable person: An egalitarian reconstruction of the objective standard*. Oxford University Press.
- Moran, M. (2010). The reasonable person: A conceptual biography in comparative perspective.
- Northrop, C. M., & Rozan, K. R. (2017). Kids will be kids: Time for a "reasonable child" standard for the proof of objective mens rea elements. *Maine Law Review*, 69(1).
- Nourse, V. (2008). After the reasonable man: Getting over the subjectivity/objectivity question. *New Criminal Law Review*, 11(1), 33–50.
- People v. White, 77 Cal. App. 3d (Cal. Ct. App. 1981).
- People v. Wu, 235 Cal. App. 3d 614 (Cal. Ct. App. 1991), ordered republished.
- Perry, E. L., Kulik, C. T., & Bourhis, A. C. (2004). The reasonable woman standard: Effects on sexual harassment court decisions. *Law and Human Behavior*, 28(1), 9–27.
- Petrocelli, J. V., Percy, E. J., Sherman, S. J., & Tormala, Z. L. (2011). Counterfactual potency. *Journal of Personality and Social Psychology*, 100(1), 30–46.
- Renteln, A. D. (2005). *The cultural defense*. Oxford University Press.
- Scalet, S. P. (2003). Fitting the people they are meant to serve: Reasonable persons in the American legal system. *Law and Philosophy*, 22(1), 75–110.
- Schultz, V. (1998). Reconceptualizing sexual harassment. *The Yale Law Journal*, 107(6), 1683–1805.
- Shaver, K. G. (1985). *The attribution of blame: Causality, responsibility, and blameworthiness*. Springer Science & Business Media.
- Shoenfelt, E. L., Maue, A. E., & Nelson, J. (2002). Reasonable person versus reasonable woman: Does it matter? *American University Journal of Gender, Social Policy & the Law*, 10(3), 633–672.
- Simpson, A., Alicke, M. D., Gordon, E., & Rose, D. (2020). The reasonably prudent person, or me? *Journal of Applied Social Psychology*, 50(5), 313–323.
- Sommers, R. (2021). Experimental jurisprudence. *Science*, 373(6553), 394–395.
- Spruill, M., & Lewis, N. A. (2023). How do people come to judge what is "reasonable"? Effects of legal and sociological systems on human psychology. *Perspectives on Psychological Science*, 18(2), 378–391.
- State v. Kelly, 97 N.J. 178 (1984).
- Stern, S. (2024). From Clapham to Salina: Locating the reasonable man. *Law & Literature*, 36(3), 391–417.
- Tobia, K. (2018). How people judge what is reasonable. *Alabama Law Review*, 70(2), 293–359.
- Tobia, K. (2022). Experimental jurisprudence. *University of Chicago Law Review*, 89(3).
- Tobia, K., Hannikainen, I. R., Kamper, D., Almeida, G., Bystranowski, P., Strohmaier, N., Dranseika, V., Kneer, M., Aguiar, F., Dolinina, K., Janik, B., Lauraitytė, E., Liefgreen, A., Próchnicki, M., Rosas, A., Kumar, V., & Struchiner, N. (2025). The nature of reasonableness. *Stanford Journal of International Law*, 61(2), 111–155.
- Votruba, A. (2013). Will the real reasonable person please stand up? Using psychology to better understand and apply the reasonable person standard. *Department of Psychology: Faculty Publications*.
- Waldmann, M. R., Nagel, J., & Wiegmann, A. (2012). Moral Judgment. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford Handbook of Thinking and Reasoning*. Oxford University Press.
- Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct*. Guilford Press.
- Westen, P. (2008). Individualizing the reasonable person in criminal law. *Criminal Law and Philosophy*, 2(2), 137–162.
- Zipursky, B. (2015). Reasonableness in and out of negligence law. *University of Pennsylvania Law Review*, 163, 2131.