Individual Linguistic Preferences Predict Liability Judgments

Parla Buyruk (pbuyruk@ucsd.edu)
Lera Boroditsky (lera@ucsd.edu)
Department of Cognitive Science, University of California, San Diego
La Jolla, CA, USA

Abstract

Can patterns in language predict people's reasoning about blame and punishment? In this study, we focus on how individual linguistic preferences (agentive vs non-agentive frames) of English, Spanish and Turkish speakers correlate with their judgments of financial liability regarding property damage across events involving intentional agents, accidental outcomes and bystanders. Findings show that individual linguistic preferences predict liability judgments, controlling for event type.

Keywords: linguistic preference**;** linguistic framing; financial liability; cross-linguistic differences; individual differences; blame; agentivity

Introduction

Can patterns in language predict people's reasoning about blame and punishment? Linguistic framing has been shown to influence how much people blame others and how liable a person is for resulting damages (Fausey & Boroditsky, 2010a; Tonković, Vlašiček & Dumančić., 2022). For example, when reports of incidents are presented using agentive frames (e.g. he "*ripped the costume*") rather than non-agentive frames (e.g. "*the costume ripped*"), people assign higher blame and financial liability, even when they see video evidence and have prior knowledge (Fausey & Boroditsky, 2010a).

Beyond framing effects induced in experiments, speakers of different languages naturally show systematic differences in how they talk about accidents. For example, Spanish and Japanese speakers use non-agentive constructions more frequently than English speakers when asked to describe simple accidents. This difference in linguistic patterns correlates with patterns in eye-witness memory for accidents (Fausey et al., 2010b; Fausey & Boroditsky, 2011).

In this paper we focus beyond linguistic habits at the language group level and ask whether specific linguistic preferences of individuals can also predict people's perceptions of agency and judgments regarding blame and punishment. If individuals, even within a language group, linguistically encode accidents differently, this might have downstream consequences in their liability judgements.

If linguistic habits indeed influence or predict patterns in thinking, then linguistic habits at the individual level should be the strongest predictors of patterns in thought. In the case of accidents, individual speakers, independent of their language background, might differ in how strongly they adopt agentive vs non-agentive frames. These individual differences in linguistic preferences may in turn predict their judgements about blame and punishment. Differences in individual linguistic preferences might predict blame and liability judgments more closely than examining linguistic preferences at the broad language group level.

The present study explores the relationship between people's linguistic preferences and their liability judgements. Beyond testing just English speakers, we also included speakers of Spanish and Turkish to broaden the examination of individual differences beyond just speakers of one language.

In the study, participants were told that one person was asking another person for money as a result of damage to their personal belongings. For each scenario, participants watched a video that looked like surveillance footage of the incident. Each participant made judgments about three events: one in which an actor performed a clearly intentional action (intentional), one where there was an accidental outcome (accidental), and one where the actor was simply a bystander to the incident and was not directly involved (bystander). Participants described what happened in each case in their own words (description task), they made judgments about the acceptability of different possible descriptions (e.g., she broke the vase vs the vase broke) (linguistic preference task), and they also made judgements about financial liability and intentionality (liability task). Half of the participants completed the description and linguistic preference tasks first and then the liability task, and the other half completed the liability task first followed by the description and linguistic preference tasks.

In this paper we focus on the results of the linguistic preference and liability judgement tasks. We ask whether individual linguistic preferences correlate with liability judgements across the three language groups included.

Methods

Participants

Participants were recruited from Prolific and completed the study online. The final sample included 276 participants, consisting of speakers of English (N = 86), Spanish (N = 93), and Turkish (N = 97). Inclusion criteria required participants to have native level proficiency in the test language. For English speakers, the sample was restricted to English-speaking monolinguals who reside in the United States. For Spanish and Turkish speakers, we screened for primary and earliest language in life to allow for a large enough participant pool.

General Design

The experiment featured a 3 (Language: English, Spanish, Turkish) × 2 (Task Order: Description & Linguistic

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

Preference task first vs Liability task first) × 3 (Event Type: Intentional, Accidental, Bystander) mixed-factorial design where Language and Task Order were between participants factors, and Event Type was within-participants. All participants completed two experimental blocks — Liability Task and Description & Linguistic Preference Task in random order.

Materials

Materials were developed in English first and then translated to Spanish and Turkish using a back-translation process and verified by independent bilingual speakers. The experiment was programmed and presented in Qualtrics.

Stimuli. A total of 9 CCTV footage-like silent videos depicting events where an object gets damaged in the presence of person(s) were created. These consisted of three different scenarios. In one a phone broke after falling from a balcony railing (Balcony), in another a phone broke after falling from a desk (Desk), and in the third a vase broke after falling off a shelf (Shelf). For each scenario we created three versions (Intentional, Accidental, Bystander) as described in detail below.

The Desk scenario features two female coworkers working at neighboring desks. The employee on the right asks the employee on the left to plug her phone to charge and she agrees. Sometime later the phone ends up on the floor and its screen gets damaged. In the intentional version, the employee on the left slides the phone off with her finger while making sure the owner of the phone isn't watching. In the accidental version, she moves her arm and the book it was resting on moves and pushes the phone accidentally off the desk. In the bystander version, a robot vacuum cleaner catches the cord and the phone falls.

The Balcony scenario features two female friends on a college campus running into each other. They chat and hug and as one of them leaves the scene she forgets to take her phone that she placed on the ledge. In the intentional version, the friend pushes the phone off the ledge while looking around to make sure nobody is watching. In the accidental version, she's resting against the ledge and as she moves she unknowingly pushes off the phone. In the bystander case the phone is seen vibrating along the ledge and falls off while she's unaware.

The Shelf scenario features a female patient in the waiting room at a dentist's office sitting in a chair next to a bookcase that has a vase next to a stack of books. In the intentional scenario she pushes over the books knocking over the vase. In the accidental version, the books fall and topple the vase off the bookcase while she's trying to take a book. In the bystander version, she's looking towards the books while they suddenly slide and fall to one side causing the vase to fall off the shelf.

All videos were 10-30 seconds long. All versions for each scenario looked visually similar and differed only on how the agent interacted with the object. In the intentional version, the agent clearly caused the damage, in the accidental version they unintentionally caused the damage, in the bystander version they were present but had no physical contact with the object when the damage occurred through other means.

Each participant saw only one version of each scenario (e.g., only the intentional of Desk, only the accidental of Balcony, only the bystander of Shelf). Each participant saw one intentional, one accidental and one bystander video and which version of which scenario was presented was rotated across participants.

Liability trials. In each trial, participants watched a video and then filled out an 8-item questionnaire. Participants were asked to make 1) a financial liability judgement on a slider question *(e.g. A new vase costs $120. The dentist who is the owner of the vase requested $120. How much should the patient pay to the owner of the vase? Hold and drag the button to select an amount between $0- $120.)* and provide written explanations for 2) why they picked this amount, 3) why not more or 4) less. Then they indicated 5) how they perceived the event *(i.e. Which of the following characterizes the event the best? The patient was… an innocent bystander/didn't mean to break it, it was an accident/broke it on purpose*) and 6) how they perceived the agent's intention to cause the damage on a 5-point scale (*i.e. Did the patient want to break the vase? Definitely yes-definitely no)*. Lastly, they rated the agent's trustworthiness on a 0-10 Likert scale on three questions (*i.e.* 6) *How do you feel towards this person? (very cold-very warm) 7) How much do you trust this person? (don't trust at all- trust completely) 8) How lucky/unlucky is this person? (very unlucky-very lucky)*.

Description & Linguistic Preference trials. In each trial, participants watched a video and answered two questions. Participants were first asked to type an open-ended response to "Please describe the event in one or two sentences. What happened?". Then they were asked "Which do you think is the most appropriate as a description of what happened?". For this question, participants selected between agentive (e.g. "she broke the vase") and non-agentive framings (e.g. "the vase broke") descriptions on a 4-point Likert scale (agentive-much more appropriate, agentive-slightly more appropriate, non-agentive slightly more appropriate, non-agentive much more appropriate).

Procedure

The study was administered online using Qualtrics. Participants first completed and IRB approved informed consent. Participants read instructions and completed all tasks in their native language. Participants were randomly assigned to complete the Description and Linguistic Preference first or the Liability trials first in random order. Each block consisted of three trials. Each trial began with an introduction to set the scene "e.g. Here you see a patient in a dentist's waiting room. After this event the dentist requested money from the patient" and participants were prompted to

watch the video. Settings ensured that participants were shown the video in full at least once, and then they were allowed to watch it again if they chose to. After they saw the video, they completed the respective tasks as described in the materials section above. After completing both tasks, they answered a demographic and language history questionnaire. The experiment took ~10-15 minutes.

Results

Overall, we found that linguistic preference was strongly associated with liability judgments, even when controlling for the type of event (intentional, accidental, bystander). Further, individual linguistic preferences were a better predictor of liability judgments than the coarser variable of Native Language group.

Across the three language groups, both linguistic preference and the liability judgments were largely driven by event type ($F(2, 819) = 799.06$, $p < .001$, and $F(2, 817) = 635.04$, $p < .001$ respectively) and no reliable cross linguistic differences emerged (Figure 1). Unsurprisingly, relative to accidental events, participants showed a stronger preference for non-agentive descriptions for bystander events and a strong preference for agentive descriptions for intentional events. Intentional events received the highest liability ratings, followed by accidental and bystander events. Critically, linguistic preferences at the individual-level were systematically related to liability judgements. Participants who showed a stronger preference for agentive frames assigned higher liability, even after controlling for event type (Figure 2). We report details of the statistical analyses below.

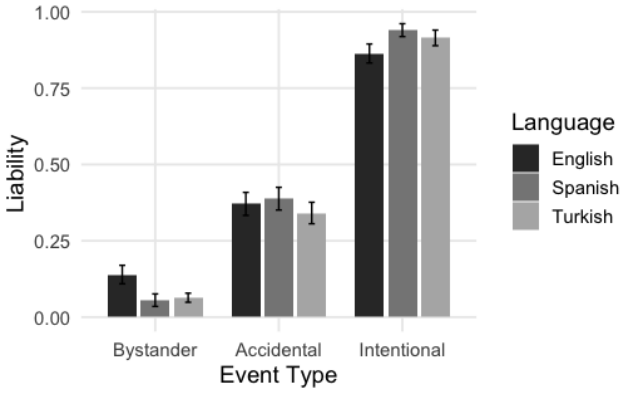

Figure 1: Mean financial liability judgements across different events in different language groups. Error bars indicate ±*1 SE*. Across all language groups, people assign higher financial liability for intentional agents compared to bystanders, while agents in accidental events receive intermediate liability.

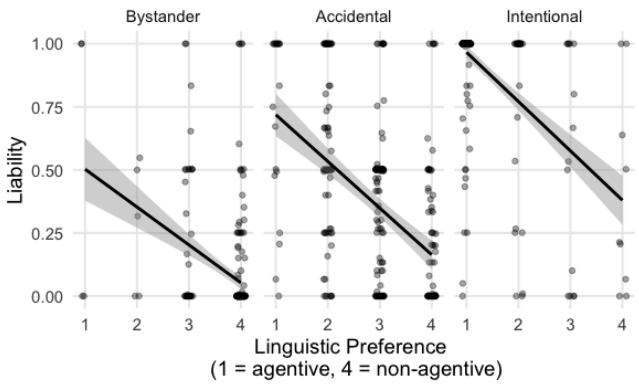

Figure 2: Relationship between linguistic preference and financial liability judgments by event type. Each panel shows a different event type (bystander, accidental, intentional). Points represent individual trial-level observations. Lines show linear fits with 95% confidence intervals. Across all event types, stronger preference for agentive descriptions (lower values on the linguistic preference scale) is associated with higher liability judgments, indicating a consistent alignment between linguistic preferences and blame attribution.

Linguistic Preferences and Liability Association. To test whether linguistic preferences predict liability judgments beyond the effect of event type, we compared a linear mixed-effects model including only Event Type to a model including both event rype and linguistic preferences. Both models had a random intercept for participants and scenarios. Adding linguistic preferences substantially improved model fit $\Delta\chi^2(1) = 223.66$, $p < .001$. While type of event alone explained 60.7% of the variance in liability judgments in the simple model, adding linguistic preferences increased explained variance to 69.6%. This indicates a strong effect of linguistic preference, as linguistic preferences uniquely explained an additional 8.9% of variance in liability and was a strong predictor for liability ($\beta = -0.18$, $SE = .01$, 95% CI [−0.20, −0.16], $t(818) = -16.07$, $p < .001$) even when controlling for Event Type. Random intercepts for participant ($SD = 0.044$) and scenario ($SD = 0.035$) were estimated from the full model.

To examine the directionality of the linguistic preference and liability association, we tested whether the linguistic preference and liability association differed depending on whether linguistic preferences were measured before or after liability judgments. The linguistic preference × task order interaction was not significant, $\Delta\chi^2(2) = 4.31$, $p = .116$, suggesting that the strength of the association did not depend on measurement order.

We next asked whether this linguistic preference–liability link differed across language groups. A model including the linguistic preference × language interaction did not improve fit compared to the simpler model, $\chi^2(4) = 4.69$, $p = .321$, and none of the interaction terms were significant. Thus, the strength and direction of the linguistic preference–liability

relationship were statistically indistinguishable across English, Spanish, and Turkish speakers.

Together, these results indicate a consistent within-participant alignment between individual linguistic preferences and liability judgments, and this relationship appears to be common across the three language groups tested.

Linguistic preference ratings. We analyzed linguistic preference using a linear mixed-effects model with event type (within-subjects) and language (between-subjects) as predictors and participant and scenario as random intercepts, with the intercept corresponding to accidental events in English speakers. This choice reflects our theoretical focus on accidental events as the ground for cross linguistic differences to emerge, and it allows model coefficients to be interpreted as deviations from this theoretical baseline. The model revealed a robust main effect of event, $F(2, 819) = 799.06$, $p < .001$. Relative to accidental events (Intercept = 2.92), participants showed a stronger preference for non-agentive descriptions for bystander events (+0.91, $p < .001$) and a strong preference for agentive descriptions for intentional events (–1.50, $p < .001$). Language did not modulate these event-type effects: the event × language interaction was not significant, $F(4, 819) = 0.71$, $p = .584$, and although a small main effect of Language emerged, $F(2, 819) = 3.23$, $p = .040$, pairwise comparisons were uniformly non-significant (all $p > .09$), indicating that the characteristic event-type gradient—bystander > accidental > intentional—was shared across language groups. Variance components for both random intercepts were negligible (SD = 0), indicating that linguistic preferences were fully explained by the fixed effects and yielding degrees of freedom equivalent to a fixed-effects model. Overall, linguistic preferences were strongly driven by event structure, with minimal cross linguistic variation. No effects of task order were observed.

Liability ratings. Participants assigned substantially different levels of liability to the three event types. A 3 (Language: English, Spanish, Turkish) × 3 (Event: Bystander, Accidental, Intentional) linear mixed-effects model with random intercepts for participant and scenario revealed a strong main effect of event, $F(2, 817) = 635.04$, $p < .001$. Intentional events received the highest liability ratings ($M = .91$, $SD = .25$), accidental events were intermediate ($M = .37$, $SD = .35$), and bystander events received the lowest liability ($M = .08$, $SD = .22$). There was no main effect of Language, $F(2, 817) = 0.50$, $p = .604$, nor a significant language × event type interaction, $F(4, 817) = 2.13$, $p = .076$. Participant variance was negligible ($SD \approx 0$); scenario variance was small ($SD = 0.047$). No effects of task order were observed. We also quantified how participants positioned accidental events on the liability continuum using the Accident Position Index (API), defined as the proportion of the intentional–bystander liability range occupied by the accidental event for each participant.

$$API = \frac{L_{accidental} - L_{bystander}}{L_{intentional} - L_{bystander}}$$

Participants whose intentional and bystander ratings were identical (yielding undefined API values) were excluded from this analysis ($n = 13$). Using a linear model with Language (between-subjects) as predictor and Task Order (between-subjects) as procedural control, we found no evidence for cross-linguistic differences in the position of accidental events, $F(2, 258) = 1.29$, $p = .28$. On average, participants placed accidental events about one-third of the way between bystander and intentional events ($M = .41$, $SD = 1.15$, range = –5.80 to 15.00), with substantial individual variability but no systematic differences across language groups. No effects involving task order were significant, nor was there a language x task order interaction.

Event comprehension checks. Participants' categorical judgments about the type of events largely matched our experimental manipulation especially for Intentional (81-89%) and Accidental (84-96%) events across all three language groups. Bystander events were more ambiguous (correct identification rate ranged 65-75%), and the misclassified cases were largely categorized as accidental. Crucially, this variability was similar across all languages, indicating that all groups understood the event structure comparably.

Intent ratings. Participants' judgments of how intentional the agent was confirmed the expected strong effect of Event type, $F(2, 817) = 1268.56$, $p < .001$, and this pattern did not differ reliably by language, $F(4, 817) = 1.61$, $p = .169$. Although a small main effect of Language emerged, $F(2, 817) = 3.01$, $p = .049$, and one pairwise comparison indicated slightly higher intent ratings for Turkish speakers on intentional trials ($t = -2.87$, $p = .012$), this effect was small and did not alter the overall pattern Participant variance was negligible ($SD \approx 0$); scenario variance was small ($SD = 0.078$). These results show that participants across languages interpreted agent's intent in highly similar ways, consistent with their construal of the events (bystander < accidental < intentional). No effects of task order were observed.

Trustworthiness ratings. Linguistic preference ratings strongly correlated with participants' evaluations of the agent where non-agentive framings accompanied higher warmth and trust ratings ($rs = .70$, $ps < .001$). Similarly, lower liability judgments strongly correlated with higher warmth and trust ratings ($rs = -.70$, $ps < .001$). Perceived luck of the agent had weak/no relationship with linguistic preference ($r = -.11$), and liability ($r = .14$).

Other measures. Open-ended responses elicited in the description task and the liability questionnaire are in the

process of being coded and therefore are not included in the paper.

Discussion

In this study, we presented participants videos of incidents resulting in property damage and manipulated agents' involvement in the outcome (Event Type: intentional, accidental, bystander). We asked participants to rate the appropriateness of agentive and non-agentive frames to measure their linguistic preferences and asked them to determine how much the agents should be responsible for damages in dollar amounts to measure their judgments of financial liability. Unsurprisingly, we found that people assigned higher liability for intentional compared to accidental and bystander events. Similarly, they showed a linguistic preference for agentive frames in intentional events and favored non-agentive frames in accidental and bystander cases. Crucially, individual-level linguistic preferences were systematically aligned with liability judgements. Those who assigned higher liability also showed a preference for agentive frames and those who assigned lower liability showed a preference for non-agentive frames, controlling for event type. Lastly, the relationship between individual linguistic preferences and liability judgments held across the three language groups (English, Spanish, and Turkish) included in our sample.

The present data provide some preliminary insights into the stability and possible cognitive basis of the linguistic preference–liability association. While the scenario-level variance suggests that the effect is partly context dependent, the consistency of the association across event types, task order, and language groups, together with participant-level variability in the models, suggests that the relationship reflects a relatively stable tendency rather than momentary task-specific responding.

The alignment between individual linguistic preferences and liability judgments is theoretically interpretable within frameworks connecting causal construal, blame attribution, and legal judgment. Agentive constructions center the agent as the causal source of an outcome, while non-agentive constructions foreground the outcome itself while backgrounding the agent's role. Research on blame attribution suggests that judgments of responsibility are shaped not only by outcomes themselves, but also by whether agents are construed as causally central to the event, capable of preventing the outcome, and in violation of normative expectations (Alicke, 1992; Malle, Guglielmo, & Monroe, 2014; Engelmann & Waldmann, 2022). Individual differences in preferred linguistic framing may therefore reflect individual differences in these underlying construal tendencies such that people who habitually center agents linguistically may also habitually construe agents as more causally responsible, which may in turn lead to higher liability judgements. Recent work in experimental jurisprudence has shown that folk causal and responsibility judgments are sensitive to factors peripheral to the legal merits of a case, including norm violations (Güver & Kneer, 2025). Our findings suggest that individual linguistic construal tendencies may represent another psychologically relevant factor for shaping responsibility and liability, with potential implications for how responsibility is assigned in real-world legal contexts such as witness testimony and jury deliberation (Tobia, 2022; Knobe & Shapiro, 2021)

A major limitation of the current study is that it is a correlational study. We measured people's linguistic preferences and liability judgments. We did not directly manipulate the linguistic framing hence making causal inferences about whether language influences liability is not possible in our design. The correlational nature of the study leaves at least three interpretations for our finding: (i) linguistic preferences influence liability judgements (ii) liability judgments influence linguistic preferences, or (iii) both may be driven by a third factor. As linguistic preference and liability association did not differ by when the linguistic preference was measured, the absence of task order interaction provides limited indirect evidence against (ii). We tentatively favor interpretation (i) in light of prior studies demonstrating that linguistic framing influences blame and liability judgments through experimental manipulation (Fausey & Boroditsky, 2010a, Tonković et al., 2022). At the same time, the present findings are also consistent with interpretation (iii), whereby both linguistic preferences and liability judgments reflect a broader tendency toward agency-centered event construal. Because linguistic preference was not experimentally manipulated in the current study, the data cannot distinguish conclusively between these possibilities.

Previous studies have experimentally manipulated linguistic framing and found effects on liability in English speakers (Fausey, 2010) and findings have been replicated by Tonković et al (2022) for English speakers but only partially for Croatian speakers. In particular, they found a smaller effect of framing on the blame judgements and no effects on financial liability. They propose that this difference reflects relative preferences for agentive frames across languages: where non-agentive descriptions are the default for accidents, agentive framing raises blame less dramatically than in languages where agentive framing is the norm. Establishing what those norms actually are through measuring linguistic preference is therefore important for understanding when and why framing effects emerge across languages.

We measured linguistic preference by asking participants acceptability of agentive vs non-agentive constructions for each event using a 4-point response scale, which did not include a neutral point. However, participants predominantly selected end points rather than the intermediate options suggesting that they held clear preferences. More importantly, while linguistic preference ratings is a proxy for event construal, a more direct measure would be how people describe these events themselves. An important next step is to compare forced-choice preferences with spontaneous event descriptions to determine whether the linguistic preference measure reflects participants'

natural production tendencies reliably. Future analyses of the elicited descriptions will provide richer data that can be used to study interindividual and cross-linguistic differences of linguistic preference. While we tested English and Spanish speakers due to established differences in how they describe events, it is worth noting that the scenarios we presented in this study are more complex and informationally rich than the simple incidents the previous studies utilized (e.g. popping a balloon). Specifically, the stimuli we created for this study featured adverse outcomes and an agent that is being held liable for essentially being purposeful or negligent in plausible real-world scenarios. Hence, analyzing people's self-generated descriptions will result in a more in-depth understanding of how people construe such events, and whether factors beyond that indexed by the linguistic preference measure in the present study would more strongly correlate with blame and liability judgements.

Historically, the linguistic relativity hypothesis has been studied cross linguistically. In contrast, Günther (2016) argues that not all individuals within a language community use the same structures with the same frequency hence focusing on individual use patterns rather than average differences across language groups is more informative for understanding language and cognition effects. In line with this view, the main finding of our study shows that speakers within the language communities we studied show variability in what they think are acceptable forms of describing events, and their individual linguistic preferences go hand in hand with their judgements of liability.

Conclusion

We found that individuals' linguistic preferences predict judgments about blame and punishment. Individuals who preferred agentive frames also assigned steeper liability judgements as compared to individuals who preferred non-agentive frames. This effect was consistent across the three language groups included in our sample. The study contributes to the understanding of the relationship between patterns in language and thought in reasoning about financial liability by focusing on patterns in individual linguistic preferences beyond coarser language-group-based analyses.

References

Alicke, M. D. (1992). Culpable causation. *Journal of Personality and Social Psychology*, *63*(3), 368.

Engelmann, N., & Waldmann, M. R. (2022). How causal structure, causal strength, and foreseeability affect moral judgments. *Cognition*, *226*, 105167.

Fausey, C. M., & Boroditsky, L. (2010a). Subtle linguistic cues influence perceived blame and financial liability. *Psychonomic Bulletin & Review*, *17*(5), 644–650.

Fausey, C. M., Long, B. L., Inamori, A., & Boroditsky, L. (2010b). Constructing Agency: The Role of Language. *Frontiers in Psychology, 1*,162.

Fausey, C. M., & Boroditsky, L. (2011). Who dunnit? Cross-linguistic differences in eye-witness memory. *Psychonomic Bulletin & Review*, *18*(1), 150–157.

Günther, F. (2016). *Constructions in Cognitive Contexts: Why Individuals Matter in Linguistic Relativity Research. In Constructions in Cognitive Contexts: Why Individuals Matter in Linguistic Relativity Research*. De Gruyter Mouton.

Güver, L., & Kneer, M. (2025). Causation, Norms, and Cognitive Bias. *Cognition*, *259*, 106105.

Knobe, J., & Shapiro, S. (2021). Proximate cause explained. *The University of Chicago Law Review*, *88*(1), 165-236.

Malle, B. F., Guglielmo, S. &. Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry*, *25*(2), 147–186.

Tonković, M., Vlašiček, D., & Dumančić, F. (2022). Preregistered direct replication of the linguistic frame effect on perceived blame and financial liability. *Legal and Criminological Psychology*, *27*(2), 354–369.