Title:

Neural evidence for narrative-based processing of evidence and bias in juror decision making.

Short Title: Narrative processing during juror decision making.

Authors:

Jaime J. Castrellon[1,2], Shabnam Hakimi[2,3], Jacob M. Parelman[3,4], Lun Yin[2], Jonathan R. Law[1,2], Jesse A.G. Skene[1,2], David A. Ball[5], Artemis Malekpour[5], Donald H. Beskind[6], Neil Vidmar[6], John M. Pearson[1,2,7,8], J. H. Pate Skene[3,8,9]*, R. McKell Carter[3,10,11]*

*** -** Senior authors contributed equally to this work.

Affiliations:

1 - Department of Psychology and Neuroscience, Duke University

2 - Center for Cognitive Neuroscience, Duke University

3 - Institute of Cognitive Science, University of Colorado Boulder

4 - Annenberg School for Communication, University of Pennsylvania

5 - Malekpour & Ball Consulting (JuryWatch, Inc.)

6 - School of Law, Duke University

7 - Departments of Biostatistics & Bioinformatics, Duke University

8 - Department of Neurobiology, Duke University

9 - Initiative in Science and Society, Duke University

10 - Department of Psychology and Neuroscience, University of Colorado Boulder

11 - Electrical, Computer, & Energy Engineering, University of Colorado Boulder

Acknowledgements: The authors would like to thank Tal Yarkoni for his assistance in accessing the Neurosynth database. This work was supported by an Incubator Award from the Duke Institute for Brain Sciences, The Institute of Cognitive Science at CU Boulder and a subcontract through eCortex, Inc. from the Office of Naval Research (ONR, N00014-18-C-2067) to R.M.C, National Science Foundation (NSF) grant no. 1655445 to J.H.P.S, and by a career development award from the NIH Big Data to Knowledge Program (grant no. K01-ES-025442 to J.M.P). J.J.C was supported by an NSF Graduate Research Fellowship (grant no. NSF DGE-1644868). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Contributions: J.H.P.S., J.M.Pe, and R.M.C. conceived the project. All authors contributed to the task design. D.A.B., A.M., D.H.B., J.A.G.S, N.V., J.M.Pe, J.H.P.S., and R.M.C. wrote the crime scenarios and evidence modules. J.M.Pa, S.H., and R.M.C designed and coded the fMRI task, and J.H.P.S. ran the fMRI experiments. S.H., J.M.Pa., and J.J.C. processed the data. J.M.Pe. and L.Y. wrote/designed the computational models of behavior. J.J.C., J.M.Pe, J.H.P.S, and R.M.C. wrote/designed the computational models of fMRI data. J.J.C. analyzed the data. J.J.C., J.H.P.S, and R.M.C. wrote the paper. All authors approved the final version of the manuscript.

Data availability:

Unthresholded fMRI statistical maps can be viewed/downloaded from Neurovault: https://neurovault.org/collections/UADNVKNI/. Code and fMRI analysis pre-registration for hypothesis grouping can be viewed and downloaded from OSF: https://osf.io/rk92x/ and Github: https://github.com/jcastrel/juror_fmri

Conflicts of interest:

The authors have no conflicts of interest to report.

Keywords:

Decision making, law, evidence, social, fMRI

Abstract: Efforts to explain complex human decisions have focused on competing models emphasizing utility, narrative, and social-affective mechanisms. These are difficult to distinguish using behavior alone. Jury decisions are among the most consequential complex decisions made in a modern society, and both narrative and utility models have been proposed to explain juror decisions. Here, we use patterns of brain activation derived from large neuroimaging databases to look for signatures of the cognitive processes associated with models of juror decision making. We asked jury-eligible subjects to rate the strength of a series of criminal cases while recording the resulting patterns of brain activation. When subjects considered evidence, cognitive processes associated with narrative models better explained the patterns of brain activation than cognitive processes associated with utility. In contrast, a biasing effect of crime type on perceived strength of the case was best explained by brain patterns associated with social cognition. Our results support a central role for narrative in integrating evidence and biases in complex decisions.

Main Text

A growing body of work in economics[1], psychology[2], and evolutionary biology[3,4] has offered narrative models as an alternative to traditional utility models as a way of understanding real-world human decisions[1,5–8]. Narrative and utility models differ significantly in the way they explain how people evaluate evidence and the influence of personal experience, emotional state, and biases in complex, real-world decisions. Utility models of decision making historically view the synthesis of evidence as primarily a quantitative process[9] and explain biases as modifications of utility calculations[10]. Alternatively, the emerging narrative framework characterizes decision making as a "goodness-of-fit" operation that depends on how well choice options match personal experience, emotional state, and biases[11]. Substantial behavioral evidence supports variants of both utility and narrative explanations of decision making. However, distinguishing behavioral predictions from the two models remains challenging. Here, we instead use neural evidence to better resolve the cognitive mechanisms supporting a complex decision.

Juror decisions about the guilt or innocence of a criminal defendant are among the most consequential decisions in society. Like other complex decisions, juror decisions require decision makers to evaluate the cumulative weight of the available evidence in light of personal experience, emotion, and biases. In contrast to fields that have relied heavily on the utility model, lawyers have largely adopted the narrative framework for trial practice[12], while psychologists and legal scholars continue to debate the role of narrative and utility models in theoretical explanations[13–16].

In the context of criminal trials, utility theory assumes that jurors assess the cumulative weight of the evidence to estimate the probability that the defendant committed the crime, and compare that probability to a threshold based on expected harms or benefits[13–15,17]. Narrative models, by contrast, propose that evidence integration relies on the extent to which the evidence can be assembled into a cohesive, compelling, and credible story[1,18]. For jury decisions, narrative explanations are based on the extent to which the total body of information about a case can be structured into a coherent account and how closely the competing narratives offered by prosecution and defense match the juror's background knowledge, experience, and beliefs[15,16,19,20].

In addition to weighing evidence directly relevant to a decision, in both the utility and narrative models, jurors may also consider factors not logically informative for deciding the facts of the case[15]. We refer to those factors as extra-evidentiary. Extra-evidentiary factors can lead to biases based on racial or cultural stereotypes or desire to convict for serious crimes that evoke strong emotional or moral responses.

Theories to explain extra-evidentiary influences outside of the utility and narrative models include: crime-type effects (generic prejudice[15,21] and leniency[22]); the liberation hypothesis[15,23]; the role of emotions[24]; moral judgment[25]; and social assessment of the future impacts of the crime[26]. These social-affective models treat extra-evidentiary information as a separate influence but an alternative is that the utility and narrative models of decision making explain both evidence and extra-evidentiary factors[13,15]. Resolving these overlapping explanations requires a new approach.

Here, we use neural data to test for the involvement of cognitive processes associated with each model. We operationalize utility, narrative, and social-affective models as sets of underlying cognitive processes. We then use these to quantify the relative contributions of these processes to complex decisions involving evidence and other factors. By relying on non-overlapping sets of cognitive processes to characterize the two models (and not any common processes like early visual cognition) we preserve the ability to weigh evidence in favor of each model. Using whole brain decoding[27], we then compare patterns of brain activation associated with those cognitive processes to those observed during the processing of evidence and extra-evidentiary information in a criminal case. Finally, we extend the whole-brain decoding approach by comparing the goodness of fit for each model of juror decision making. This new approach allows us to resolve the contributions of each psychological model to a given decision.

Results

To implement this approach, we collected functional magnetic resonance imaging (fMRI) data from jury-eligible participants as they viewed a series of simplified crime scenarios that vary in the type of crime and amount of evidence supporting guilt[28]. Using this task (Fig. 1A), we previously showed that evidence and extra-evidentiary information about the crime independently influence participant ratings of the strength of the case against the accused[28]. The sum of these ratings tracks the accumulation of evidence (Fig. 1B). Consistent with the previous results, computational modeling of ratings by the current participants also distinguished the effects of different types of evidence and types of crime on case strength (Fig. 1C). The extra-evidentiary influence of each type of crime on case strength is correlated with the severity of the crime, as measured by ratings of deserved punishment (Fig. S3). Textual analysis indicates that the loading of each scenario along the dimension defined by this correlation is driven by terms describing the type of crime, rather than other attributes of the scenario text (Fig. S4). We therefore refer to this effect on the perceived strength of the case strength as crime-type bias. It captures the extent to which the seriousness of a crime increases the rated strength of the case, independent of the evidence. This is akin to generic prejudice[15,21].

To identify the patterns of brain activation associated with evidence accumulation and crime-type bias, we modeled the neuroimaging data using separate parametric regressors for each (see SI). We observed distinct constellations of activations associated with evidence accumulation (Fig. 2A, SI) and crime-type bias (Fig. 2B, SI). These constellations consist of multiple regions, each associated with separate cognitive processes. In order to evaluate support for different models of decision making, we compared the constellations of activations for evidence and crime-type bias to patterns taken from a meta-analytic database, Neurosynth [29]. This meta-analytic database identifies brain activation patterns associated with defined terms and topics from the neuroimaging literature. From a set of 200 topics, we selected groups to form non-overlapping sets of cognitive processes associated with utility, narrative, and social-affective models of decision making (Table S1). We then compared thresholded brain-activation maps from the evidence accumulation and crime-type bias regressors to the cognitive processes associated with each psychological model of juror decision making.

We preregistered the list of topics associated with each model and submodel (https://osf.io/rk92x/) to limit user degrees of freedom while defining model structure. For the utility model, we included core cognitive topics associated with value, the consideration of risk, and mathematical calculation. For the narrative model, we included a series of submodels: 1. Experience, culture, and recall narrative (meant to encompass models associated with the consistency of a story and past experiences). This model was further subdivided into: 1a. culture and ideation bias and 1b. recall; 2. Working memory narrative (meant to capture the process of holding the pieces of a story together); and 3. Reading narrative (meant to capture models associated with processing fluency). For the social-affective model, we considered both affect and moral judgment. Within moral judgment, we included submodels for the process of applying morality and broader studies of social cognition. Hereafter, a model is described as significant if it explains more variance than 95% of models composed of randomly drawn topics from the 177 topics not associated with one of the a priori models (See SI). The null distribution for each test was defined for a number of topics matched to each model (e.g., 5 drawn topics if the model being evaluated contained 5 topics)

Evidence accumulation

We next describe how well each model is correlated with the patterns of brain activation observed for evidence accumulation (Fig. 2A) and crime-type bias (Fig. 2B). If evidence accumulation is largely a matter of quantitative reasoning regarding the likelihood that an event occurred, we would expect that patterns of brain activation associated with evidence accumulation would be similar to topic maps from the utility

model. On the other hand, if evidence accumulation relies on non-quantitative processes, like the coherence and plausibility of a narrative, then we would expect brain activation for evidence accumulation to correspond to topic maps from the narrative models. For example, the recall of past experiences and events is more likely to be necessary when evaluating narrative plausibility. Patterns of brain activation during evidence accumulation were significantly correlated with topic maps from the narrative model ($R^2 = 0.090$, $p = 0.014$, resampled null, Fig. 3A, see SI for details) but not utility ($R^2 = 0.046$, $p = 0.24$, resampled null, Fig. 3A) or social-affective ($R^2 = 0.033$, $p = 0.72$, resampled null, Fig. 3A) topic sets. Of the narrative submodels, both the recall narrative and reading narrative topic maps explained a significant percentage of the variance (Fig. 3A, $R^2 = 0.042$, $p = 0.019$, resampled nul and $R^2 = 0.065$, $p = 0.020$, resampled nul, respectively). The reading topic model includes cognitive processes associated with both the reading of individual words and sentences and the integration of words and sentences to interpret the meaning of the text. In our analysis, the activation of brain regions associated with the reading submodel correlated parametrically with the cumulative weight of the evidence (Fig. 3A), and not with the cumulative number of words or complexity/readability in the evidence text (Fig. S8). The correlation between patterns of brain activation observed with increasing evidence and patterns associated with recall provides support for the role of narrative in weighing evidentiary contributions to case strength.

Crime-type bias

We next asked which of our models correlate with patterns of brain activation for a specific type of extra-evidentiary information, crime-type bias. To do so, we compared the patterns of brain activation for crime-type bias to the patterns from the topics assigned to each model. In this comparison, patterns of brain activation for crime-type bias corresponding to the topic maps for either utility or narrative would support a direct influence of extra-evidentiary information on utility or narrative processes. Alternatively or in addition, social-affective topic maps could also produce patterns of brain activation associated with crime-type bias independently or interacting with utility[30] or narrative models. We find that patterns of brain activation in response to crime-type bias (Fig 2B) were strongly correlated with only two submodels, the social (theory of mind) component of the moral judgment submodel ($R^2 = 0.045$, $p = 0.012$, resampled null, Fig. 3B), and the culture and ideation bias topic maps from the experience, culture, and recall submodel ($R^2=0.054$, $p=0.0016$, resampled null, Fig. 3B). No other submodel or higher-level model explained significant variance (Fig. 3B, SI). Furthermore, the pattern of activation associated with culture and ideation bias is nearly identical to the social cognition topic from the social-affective model (Fig. 4A). This indicates that crime-type

bias in our results can be explained by a single cognitive process, social cognition. Consistent with this conclusion, we found that those scenarios that result in bodily harm or death (but not victimless or property loss crimes) drove both the increases in case-strength ratings (Fig. 4B, SI) and neural activation associated with crime-type bias (Fig. 4C, SI). Thus, we found that crime-type bias was associated with activation of brain regions and cognitive processes distinct from those associated with evidence accumulation (compare Fig. 3A to 3B).

Model-independent verification of narrative role in juror decisions

Because the particular cognitive processes assigned to each model of juror decision making are subjective, we employed a data-driven approach to ask which of all 200 possible topics were most strongly associated with evidence accumulation and crime-type bias. To do so, we compared the patterns of activation from evidence accumulation and crime-type bias to reverse-inference maps for all 200 topic maps included in the Neurosynth database. We calculated the Pearson correlation between the thresholded parametric evidence accumulation map and the parametric crime-type-bias map and the 200 topic maps in Neurosynth [29] (see SI). These findings confirmed, and further detailed, the role of narrative recall and mathematical calculation in evidence accumulation but also added a separate topic (“reasoning”) we had not included in our a priori models (Fig. 5A).

Topics with the strongest correlations for crime-type bias were centered on social cognition, directly (“mind, mental, social”, “social, interactions”, “type, joint”), indirectly (“gestures, abstract, race”, “body, bodies”), or anatomically through regions associated with social cognition (“junction, tpj, temporoparietal”, “temporal, sulcus, superior”) (Fig. 5B). Our a priori models assigned the two topics most correlated with crime-type bias (“gestures, abstract, race” and “mind, mental, social”) to experience, culture and recall and moral judgment respectively. This provides additional support for the central role of social cognition in weighing extra-evidentiary information.

Discussion

We have used neuroimaging to quantify the contributions of utility and narrative models to jury decision making, an example of complex human decisions. To do that, we extended a novel method for inferring cognitive processes from patterns of brain activation. Together, our findings suggest a prominent role for the narrative model in juror decision making that could be extended to other complex decisions. Both utility and narrative models recognize that human decisions incorporate both evidence and

extra-evidentiary influences, including personal experience, emotional states, and biases. However, narrative and utility frameworks hypothesize different mechanisms by which evidence and extra-evidentiary factors contribute to a decision. Our results indicate that cognitive mechanisms underlying the narrative model explain more of the variation in perceived case strength due to both evidence accumulation and crime-type bias. Thus, the cognitive processes engaged during juror decision making provide support for the narrative model of complex decision making.

Recall and reading processes track evidence accumulation

In addition to providing evidence of the importance of the narrative model, we were able to identify specific cognitive processes from the narrative model that significantly contributed to evidence accumulation in this task. We note that our analysis examined only the pattern of activity positively correlated evidence accumulation and does not address the negative correlated regions (Fig. 2A, though see SI). Recall of past experiences[31] and reading are the only significant submodels that contribute to the explained variance. Recall allows the comparison of the current evidence to past experiences, stories, or schema[31], which can serve as a basis for assessing the plausibility and cumulative weight of the evidence. Involvement of the reading submodel in evidence accumulation likely reflects the cohesion of the individual items of evidence rather than lower-level reading processes. The reading submodel includes not only the processes associated with reading individual words and sentences, but also the cognitive processes involved in integrating the content of words and sentences over time to assign meaning to the text. Consistent with this interpretation, imaging studies of subjects listening to stories have found that the integration of stories over increasing temporal windows shifts the pattern of brain activation from areas near auditory cortex to brain regions overlapping with those in our reading submodel[32].

Previous work has shown that the familiarity and ease of processing a particular idea, including how easily it is read, is strongly associated with confidence in the validity of a statement[33–35]. Mentally, reading engagement might serve as an index of familiarity or the ease of understanding evidence that comprises a coherent account, and therefore confidence in the story presented, as proposed in the Story Model of juror decisions[18]. Neuroimaging work comparing brain activation between true and false sentences generally supports this account, finding greater activation in the temporal parietal junction for true compared to false statements[36]. We note that our findings cannot exclude a role for quantitative reasoning in evidence accumulation. In fact, while the set of topics included in our utility model do not meet the statistical threshold, our model-independent analysis indicates that arithmetic/quantitative terms are ranked third in terms of the variance in evidence accumulation explained (Fig. 5A). Overall, however,

our results do not provide significant support for quantitative processes as the primary cognitive basis for assessing the cumulative weight of the evidence but are consistent with hybrid models in which narrative recall serves as a basis for evaluating evidence within a utility framework[37].

Social cognitive processes track crime-type bias

We found previously that the type of crime with which a defendant is charged influences the perceived strength of the case against a defendant, independent of the evidence[28]. This crime-type bias, replicated here (Fig. S3), provides an opportunity to examine the cognitive processes by which extra-evidentiary factors influence complex decisions. For crime-type bias, as for evidence accumulation, the higher-level narrative model again explains almost as much variance as it does above for evidence accumulation. However, the results for crime-type bias differ from evidence accumulation in two important ways. First, within the narrative model, the submodel that drives the contribution of narrative to crime-type bias is culture and ideation bias, while for evidence accumulation it was recall and reading. Second, in contrast to evidence accumulation, social-cognitive processes play a significant role in crime-type bias.

Although culture and ideation bias and social cognition appear as separate processes in our models, the patterns of activation associated with those two submodels are nearly identical (Fig. 4A), reflecting a common underlying use of social cognition. While our a priori models did not include social cognition as a core component of narrative stories, it has been incorporated into broader models of narrative decision making[1]. In fact, the contagious nature of a narrative has been tied to the role social value plays in making decisions about others[38]. Others have suggested that storytelling has played an important role in the evolution of cooperation in human societies by serving as means of transmitting cultural beliefs and knowledge beyond the first-hand experience of any individual[39–41]. Within this framework, social cognition generally (and culture and ideation biases in particular) could be understood as providing a basis for evaluating the plausibility of a narrative suggested by uncertain or incomplete evidence in a particular case. We found that increased social cognitive processing was dependent on the presence of bodily harm. Consistent with that finding, recent work has shown that attempts to bridge individuals with disparate beliefs are more likely to succeed when they include stories that relate personal experiences of harm[42]. More broadly, others have noted the overlap of brain regions involved in narrative and social cognition[43]. Our results are therefore consistent with crime-type bias driven by social cognition as a component of narrative processing.

While we have examined one specific type of bias that applies specifically to juror decisions in criminal justice, the cognitive model we describe here suggests a more

general context for understanding the interaction of biases, stereotypes, and cultural prejudices in complex decisions. In this view, bias does not require emotion or judgment, a view consistent with connections between social neuroscience and bias proposed by others[44]. The overlap between neural correlates of social cognition and bias (noted also by others[45,46]) implies that biases, stereotypes, and cultural beliefs are an important part of the knowledge base by which people make predictions about the world and evaluate alternative explanations. The incorporation of social cognition in the narrative model provides a means by which bias (how likely is this story to happen in general?) and evidence (what is the magnitude of support for this story?) can be integrated to reach a conclusion.

Relationship between crime-type bias and third-party punishment

Notably, neither moral judgment[47] nor affect[48] contributed significantly to crime-type bias. This is initially surprising because crime-type bias, as we operationalized it here, is closely correlated with subject ratings of the punishment deserved by each type of crime (Fig. S3). Previous studies have found that third-party punishment is correlated with activation of brain regions associated with affect and norm enforcement, in some instances modulated by the temporal parietal junction and other regions associated with social cognition[26,49–55].

These differences emphasize that crime-type bias and punishment decisions represent separable processes. Punishment ratings reflect a direct assessment of crime seriousness, while crime-type bias represents an indirect effect of crime seriousness on a decision about who committed the crime. Moral judgment has been associated with the temporal parietal junction, but the topic maps associated with moral judgment also include areas associated with motivation (e.g., the basal ganglia, topic 135) that were not present in the crime-type bias activations. Similarly, for affect, brain regions involved in affective responses (e.g., amygdala) have been linked to moral judgments about actions that harm others [49,56] but our results show no direct link between affect and crime-type bias in decisions about who committed a crime.

In that regard, it is important to point out that our experimental task, by design, presents only a minimal description of the crime and the evidence for each scenario, without graphic details or discussion of the harm to victims that might increase emotional or moral responses. Nonetheless, our modeling of the behavioral results shows that these minimalist crime descriptions are sufficient to elicit a robust bias towards greater confidence in the guilt of someone accused of more serious crimes (Fig. 1C). We therefore conclude that a robust crime-type bias can persist even in the absence of factors that trigger strong affective responses or moral judgments, and that bias under those conditions can be driven by social cognitive processes (Fig. 3B).

Conclusion

Our results thus support a model in which evidence accumulation and crime-type bias influence complex decisions through distinct, but complementary, cognitive processes. A narrative framework provides one coherent model for the integration of these two pathways in a decision about the strength of the case against the defendant in a particular case. In that model, an assessment of the cumulative weight of the evidence depends on the ease (fluency) of assembling the individual items of evidence into a coherent narrative that is credible in light of the juror's past experience and knowledge. Crime-type bias reflects a further source of information that jurors can use to assess the credibility of the prosecution's narrative, based on the social understanding of other people's behavior, including biases, stereotypes, and cultural beliefs. This offers a process by which biases can influence complex decisions even in the absence of strong affect.

In addition to informing best practices in the legal system, we offer this study as an example of a brain-decoding approach to elucidate a broad range of complex decision making processes in humans. Recent work has emphasized the role of social cognition and story-telling in human evolution[4]. And we note that the brain regions associated with both evidence accumulation and extra-evidentiary influences on complex decision making are highly expanded in modern humans compared to non-human primates[57–59] and Neanderthals[60–63]. This suggests that similar brain regions and their associated cognitive processes could be involved in a broad range of complex human decisions. In particular, narrative and social cognitive processes linked to juror decision making may offer a template for integrating evidence and social/cultural influences in other consequential decisions[64]. The approach we describe here might therefore be applied to quantifying the relative contributions of narrative, utility, and extra-evidentiary factors in a range of economic and social phenomena from asset pricing to the spread of fake news and market bubbles.

Fig. 1. A mock-juror task separates evidence accumulation and crime-type bias. During the task (**A**), participants read a criminal scenario paired with variable evidence from each of three types and rated the strength of the case and the recommended punishment severity. (**B**) Case-strength contributions from scenario independent of evidence and case-strength contributions from evidence independent of scenario. Symbols represent mean effect size (scale 0-100); error bars represent 95% credible intervals. Scenario and evidence types depicted in panel B are distinguished with triangles. Scenario effects are shown in rank order. (**C**) An example combination of scenario and evidence types is plotted as individual filled triangles for the accumulation of mean case-strength ratings (scale 0-100).

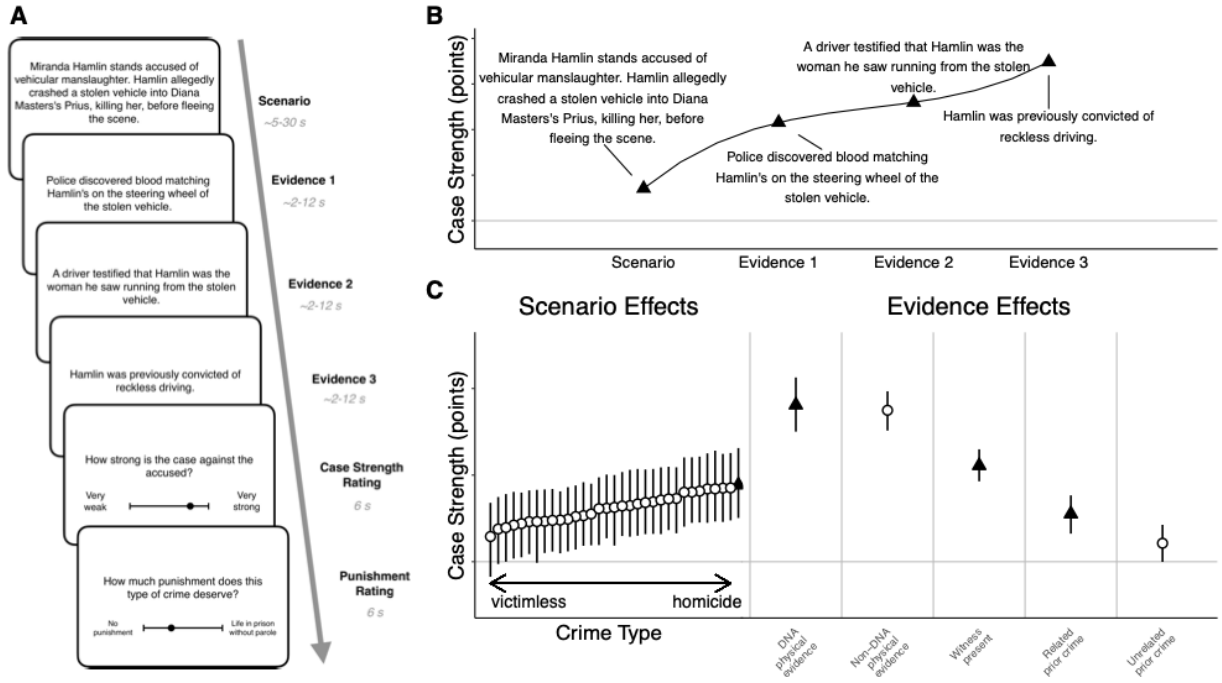


Fig. 2. Accumulation of evidence and crime-type bias produce different fMRI activation patterns. FMRI results are shown for (**A**) evidence accumulation during evidence presentation and (**B**) crime-type bias during scenario presentation. Results shown are corrected for multiple comparisons using a whole brain cluster-forming threshold Z > 2.3, cluster-corrected $p < 0.05$.

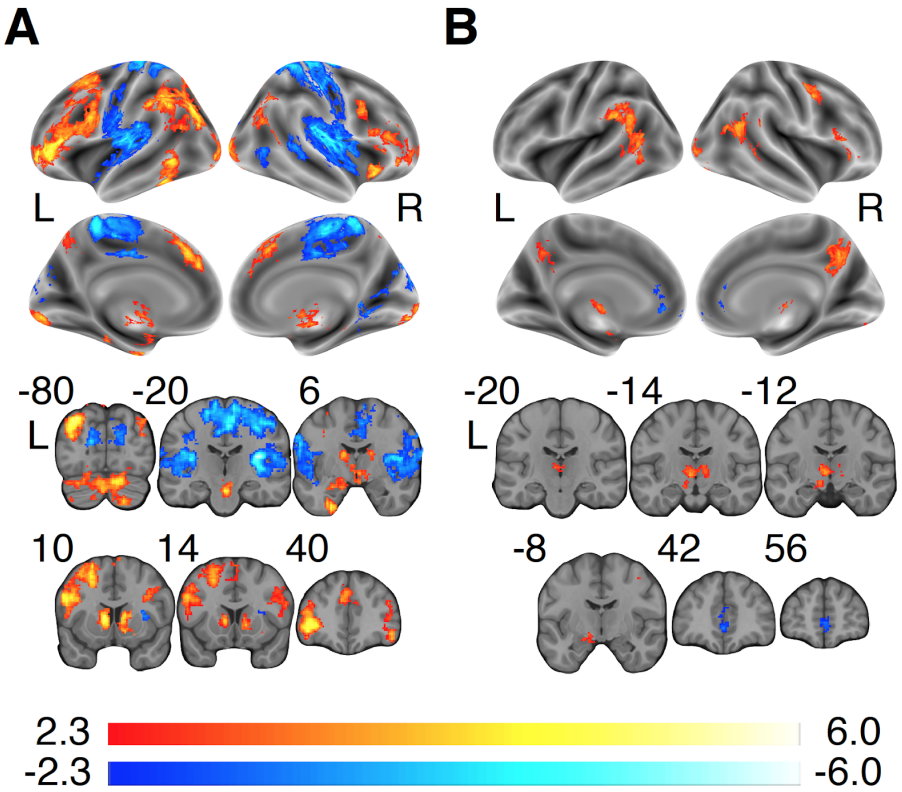


Fig. 3. Patterns of brain activation associated with evidence accumulation and crime-type bias support different decision-making models. Height of dendrograms shows the similarity (R^2) of thresholded parametric fMRI maps with Neurosynth topic models and submodels for (**A**) evidence accumulation and (**B**) crime-type bias. Stars indicate models with R^2 values that are higher than the 95 percentile of equally-sized random models (see SI).

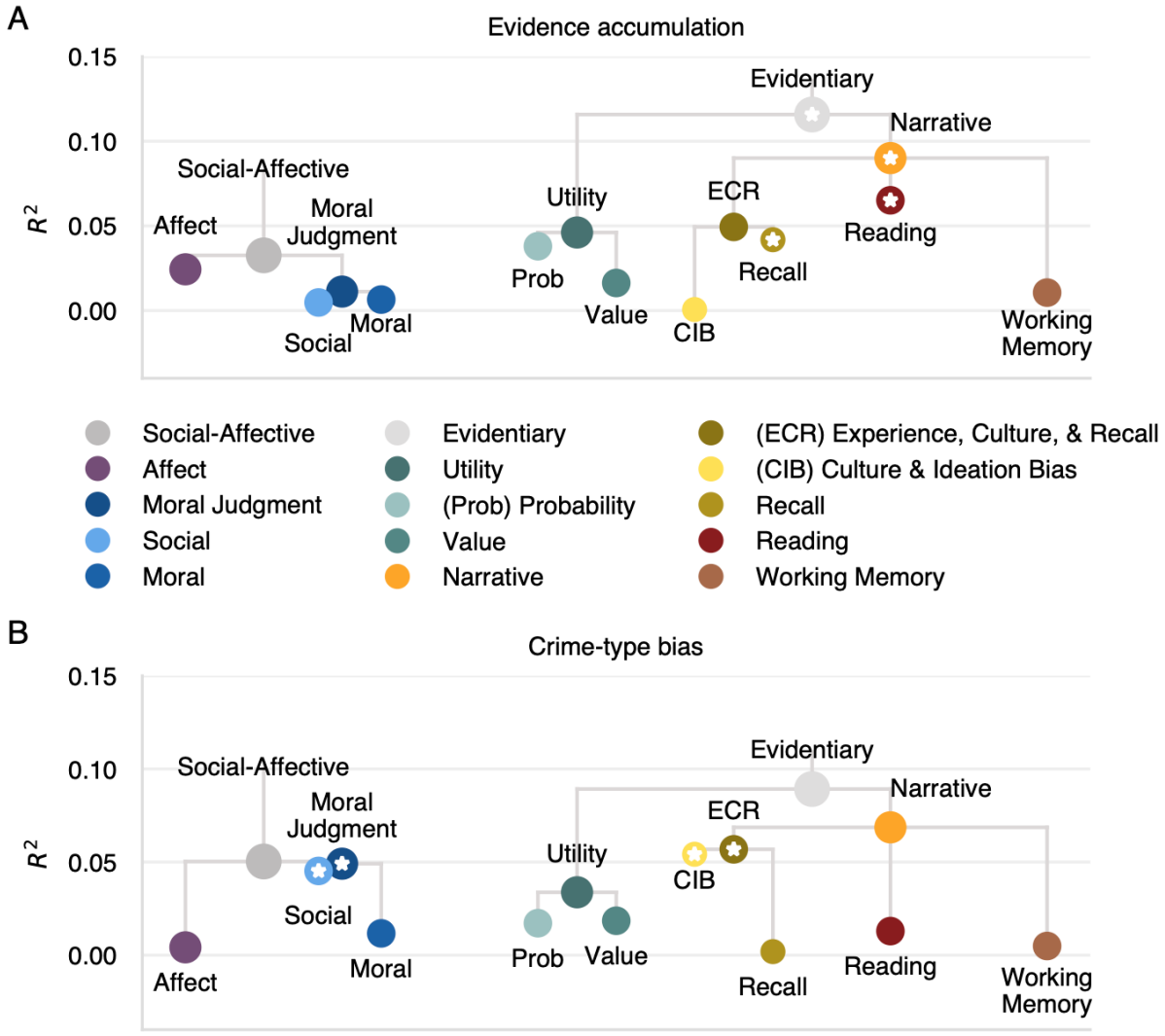


Fig. 4. Social cognitive processes explain crime-type bias. (**A**) Neurosynth reverse-inference maps from topics defined as belonging to culture and ideation bias (topic map 30: “events, future, personal” and topic map 100: “gestures, abstract, race”) and social cognition (topic map 145: “mind, mental, social” and topic map 154: “social, interactions”) decision making models. Culture and ideation bias maps are summed and shown in red, social cognition maps are summed and shown in blue, and their overlap is shown in purple. Maps are thresholded to correct for multiple comparisons (false-discovery rate threshold = 0.01). (**B**) Model estimates of crime-type bias from ratings (PC1) and (**C**) mean fMRI activation from patterns associated with crime-type bias depend on the presence of an identifiable victim. Crimes that resulted in harm to a person were uniquely associated with higher cognitive ($F(2,30) = 32.08$, $p < 0.001$) and neural ($F(2,926) = 13.04$, $p < 0.001$) measures of crime-type bias than crimes without a victim or crimes that resulted in a loss of property. Error bars represent ± 1 standard error.

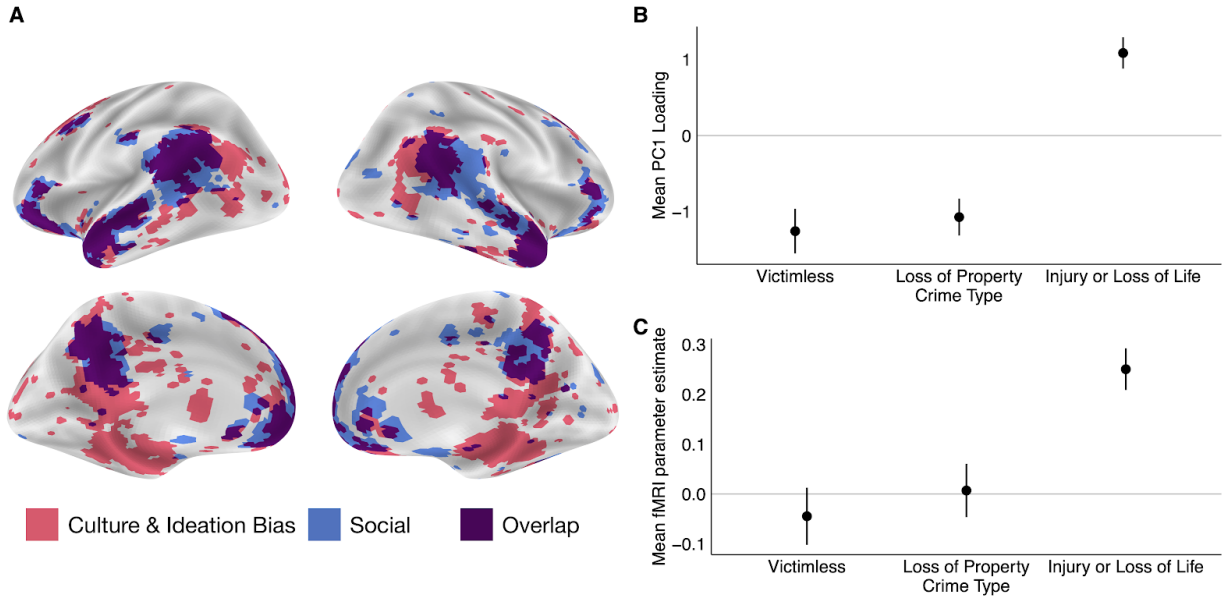
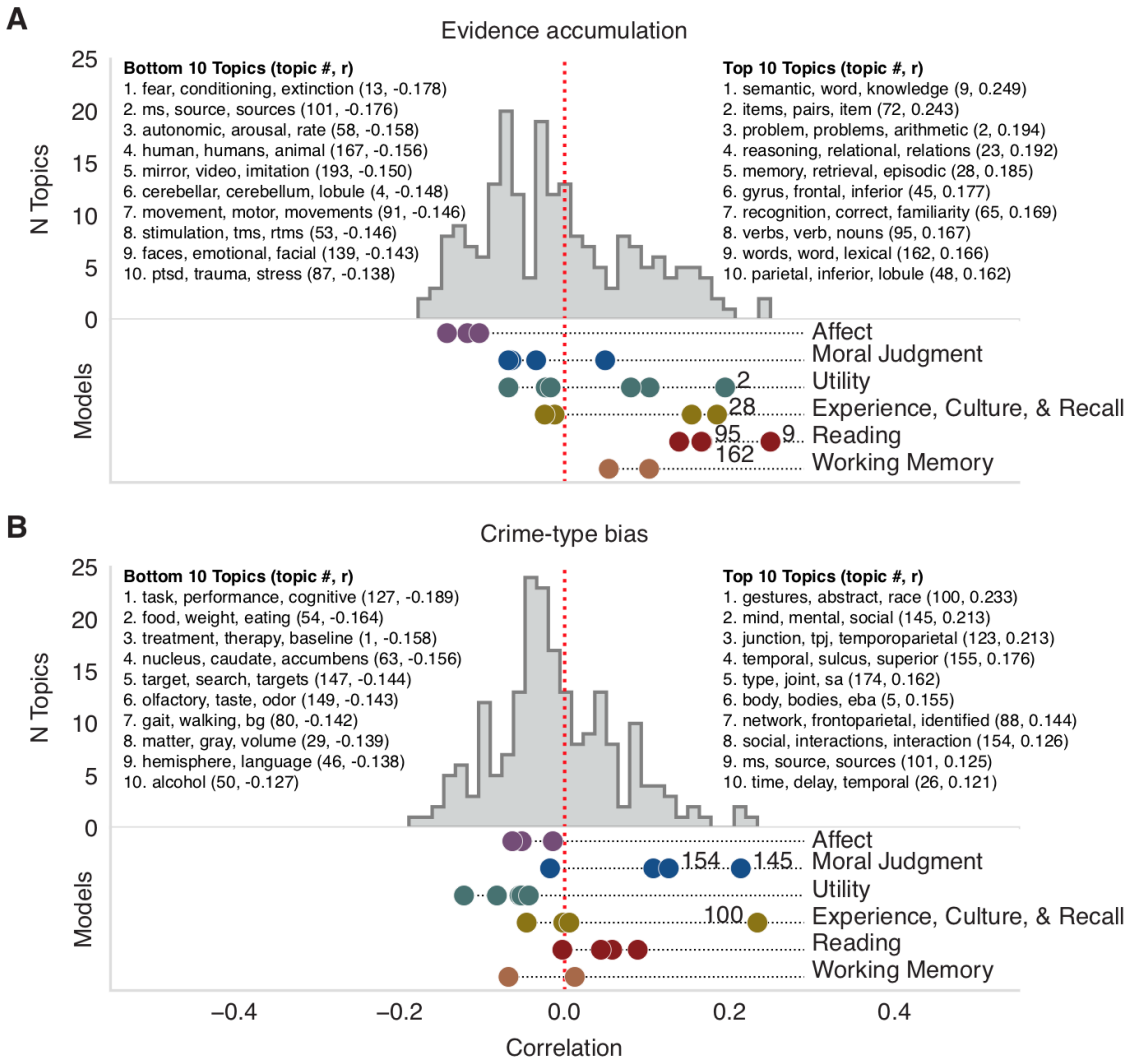


Fig. 5. Brain activation patterns associated with evidence accumulation and crime-type bias are associated with different cognitive processes. Histograms represent the distribution of spatial correlations between 200 Neurosynth topic maps and fMRI maps for (**A**) evidence accumulation and (**B**) crime-type bias. Below each histogram are points that correspond to the location of the correlation coefficients in the histogram for each topic included in each of the indicated decision-making models. The top and bottom 10 correlated topics from the entire set of 200 topics are indicated beside the histograms (top panels) and topics within each model that appear in the top 10 list are labeled beside their associated point for each model (bottom panels).



Bibliography

1. Shiller, R. J. Narrative Economics. *Am. Econ. Rev.* **107**, 967–1004 (2017).
2. Tuckett, D. & Nikolic, M. The role of conviction and narrative in decision-making under radical uncertainty. *Theory Psychol.* **27**, 501–523 (2017).
3. Gottschall, J., Wilson, D. S. & Wilson, E. O. *The Literary Animal: Evolution and the Nature of Narrative*. (Northwestern University Press, 2005).
4. Wiessner, P. W. Embers of society: Firelight talk among the Ju/'hoansi Bushmen. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 14027–14035 (2014).
5. Cialdini, R. B. & Goldstein, N. J. Social influence: compliance and conformity. *Annu. Rev. Psychol.* **55**, 591–621 (2004).
6. Tremblay, S., Sharika, K. M. & Platt, M. L. Social Decision-Making and the Brain: A Comparative Perspective. *Trends Cogn. Sci.* **21**, 265–276 (2017).
7. Thaler, R. H. & Sunstein, C. R. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. (Penguin, 2009).
8. Menon, A. Bringing cognition into strategic interactions: S trategic mental models and open questions. *Strategic Manage. J.* **39**, 168–192 (2018).
9. von Neumann, J. & Morgenstern, O. *Theory of Games and Economic Behavior*. 776 (Princeton University Press, 1944).
10. Kahneman, D. Maps of bounded rationality: Psychology for behavioral economics. *Am. Econ. Rev.* **93**, 1449 (2003).
11. Arrow, K. J. Utilities, Attitudes, Choices: A Review Note. *Econometrica* **26**, 1–23 (1958).

12. Lubet, S. & Lore, J. C. *Modern Trial Advocacy Analysis & Practice: Fifth Edition (NITA)*. (Wolters Kluwer, 2015).
13. Connolly, T. Decision theory, reasonable doubt, and the utility of erroneous acquittals. *Law Hum. Behav.* **11**, 101–112 (1987).
14. Arkes, H. R., Shoots-Reinhard, B. & Mayes, R. S. Disjunction Between Probability and Verdict in Juror Decision Making. *J. Behav. Decis. Mak.* **25**, 276–294 (2012).
15. Devine, D. J. *Jury Decision Making: The State of the Science*. (NYU Press, 2012).
16. Hastie, R. The case for relative plausibility theory: Promising, but insufficient. *The International Journal of Evidence & Proof* **23**, 134–140 (2019).
17. Kaplan, J. Decision Theory and the Factfinding Process. *Stanford Law Rev.* **20**, 1065–1092 (1968).
18. Pennington, N. & Hastie, R. Evidence evaluation in complex decision making. *J. Pers. Soc. Psychol.* **51**, 242–258 (1986).
19. Allen, R. J. & Pardo, M. S. Relative plausibility and its critics. *The International Journal of Evidence & Proof* **23**, 5–59 (2019).
20. Arkes, H. R. & Garske, J. P. *Psychological Theories of Motivation*. (Brooks/Cole, 1982).
21. Vidmar, N. Case studies of pre- and midtrial prejudice in criminal and civil litigation. *Law Hum. Behav.* **26**, 73–105 (2002).
22. Kerr, N. L. Severity of prescribed penalty and mock jurors' verdicts. *J. Pers. Soc. Psychol.* **36**, 1431–1442 (1978).
23. Kalven, H., Zeisel, H., Callahan, T. & Ennis, P. *The american jury*. (Little, Brown

Boston, 1966).

24. Holloway, C. & Wiener, R. L. The Role of Emotion and Motivation in Jury Decision-Making. *Criminal Juries in the 21st Century: Psychological Science and the Law* 247 (2018).

25. Greene, J. & Haidt, J. How (and where) does moral judgment work? *Trends Cogn. Sci.* **6**, 517–523 (2002).

26. Buckholtz, J. W. & Marois, R. The roots of modern justice: cognitive and neural foundations of social norms and their enforcement. *Nat. Neurosci.* **15**, 655–661 (2012).

27. Li, R. *et al.* Reason's Enemy Is Not Emotion: Engagement of Cognitive Control Networks Explains Biases in Gain/Loss Framing. *J. Neurosci.* **37**, 3588–3598 (2017).

28. Pearson, J. M. *et al.* Modelling the effects of crime type and evidence on judgments about guilt. *Nat Hum Behav* **2**, 856–866 (2018).

29. Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C. & Wager, T. D. Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* **8**, 665–670 (2011).

30. Roberts, I. D. & Hutcherson, C. A. Affect and Decision Making: Insights and Predictions from Computational Models. *Trends Cogn. Sci.* **23**, 602–614 (2019).

31. Baldassano, C., Hasson, U. & Norman, K. A. Representation of Real-World Event Schemas during Narrative Perception. *J. Neurosci.* **38**, 9689–9699 (2018).

32. Lerner, Y., Honey, C. J., Silbert, L. J. & Hasson, U. Topographic Mapping of a

Hierarchy of Temporal Receptive Windows Using a Narrated Story. *Journal of Neuroscience* **31**, 2906–2915 (2011).

33. Hasher, L., Goldstein, D. & Toppino, T. Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior* **16**, 107–112 (1977).
34. Arkes, H. R., Hackett, C. & Boehm, L. The generality of the relation between familiarity and judged validity. *J. Behav. Decis. Mak.* **2**, 81–94 (1989).
35. Fazio, L. K., Brashier, N. M., Payne, B. K. & Marsh, E. J. Knowledge does not protect against illusory truth. *J. Exp. Psychol. Gen.* **144**, 993–1002 (2015).
36. Marques, J. F., Canessa, N. & Cappa, S. Neural differences in the processing of true and false sentences: insights into the nature of 'truth' in language comprehension. *Cortex* **45**, 759–768 (2009).
37. Shadlen, M. N. & Shohamy, D. Decision Making and Sequential Sampling from Memory. *Neuron* **90**, 927–939 (2016).
38. Scholz, C. *et al.* A neural model of valuation and information virality. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 2881–2886 (2017).
39. Smith, D. *et al.* Cooperation and the evolution of hunter-gatherer storytelling. *Nat. Commun.* **8**, 1853 (2017).
40. Bietti, L. M., Tilston, O. & Bangerter, A. Storytelling as Adaptive Collective Sensemaking. *Top. Cogn. Sci.* **11**, 710–732 (2019).
41. Hitchcock, R. K. Hunters and gatherers past and present: Perspectives on diversity, teaching, and information transmission. *Rev. Anthr.* **48**, 5–37 (2019).
42. Kubin, E., Puryear, C., Schein, C. & Gray, K. Personal experiences bridge moral

and political divides better than facts. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).

43. Yuan, Y., Major-Girardin, J. & Brown, S. Storytelling Is Intrinsically Mentalistic: A Functional Magnetic Resonance Imaging Study of Narrative Production across Modalities. *J. Cogn. Neurosci.* **30**, 1298–1314 (2018).
44. Amodio, D. & Cikara, M. The Social Neuroscience of Prejudice. (2020) doi:10.31234/osf.io/2sdzb.
45. Harris, L. T. & Fiske, S. T. Social neuroscience evidence for dehumanised perception. *European Review of Social Psychology* **20**, 192–231 (2009).
46. Cikara, M. & Van Bavel, J. J. The Neuroscience of Intergroup Relations: An Integrative Review. *Perspect. Psychol. Sci.* **9**, 245–274 (2014).
47. Young, L., Cushman, F., Hauser, M. & Saxe, R. The neural basis of the interaction between theory of mind and moral judgment. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 8235–8240 (2007).
48. Kragel, P. A. & LaBar, K. S. Decoding the Nature of Emotion in the Brain. *Trends Cogn. Sci.* **20**, 444–455 (2016).
49. Treadway, M. T. *et al.* Corticolimbic gating of emotion-driven punishment. *Nat. Neurosci.* **17**, 1270–1275 (2014).
50. Ginther, M. R. *et al.* Parsing the Behavioral and Brain Mechanisms of Third-Party Punishment. *J. Neurosci.* **36**, 9420–9434 (2016).
51. Stallen, M. *et al.* Neurobiological Mechanisms of Responding to Injustice. *J. Neurosci.* (2018) doi:10.1523/JNEUROSCI.1242-17.2018.
52. Buckholtz, J. W. *et al.* The neural correlates of third-party punishment. *Neuron* **60**,

575 930–940 (2008).

576 53. Zinchenko, O. Brain responses to social punishment: a meta-analysis. *Sci. Rep.* **9**,
577 12800 (2019).

578 54. De Quervain, D. J. *et al.* The neural basis of altruistic punishment. *Science* **305**,
579 1254–1258 (2004).

580 55. FeldmanHall, O. *et al.* Differential neural circuitry and self-interest in real vs
581 hypothetical moral decisions. *Soc. Cogn. Affect. Neurosci.* **7**, 743–751 (2012).

582 56. Ngo, L. *et al.* Two Distinct Moral Mechanisms for Ascribing and Denying
583 Intentionality. *Sci. Rep.* **5**, 17390 (2015).

584 57. Croxson, P. L., Forkel, S. J., Cerliani, L. & Thiebaut de Schotten, M. Structural
585 Variability Across the Primate Brain: A Cross-Species Comparison. *Cereb. Cortex*
586 **28**, 3829–3841 (2018).

587 58. Smaers, J. B. & Vanier, D. R. Brain size expansion in primates and humans is
588 explained by a selective modular expansion of the cortico-cerebellar system.
589 *Cortex* **118**, 292–305 (2019).

590 59. Van Essen, D. C. *et al.* Cerebral cortical folding, parcellation, and connectivity in
591 humans, nonhuman primates, and mice. *Proc. Natl. Acad. Sci. U. S. A.* (2019)
592 doi:10.1073/pnas.1902299116.

593 60. Bruner, E., Preuss, T. M., Chen, X. & Rilling, J. K. Evidence for expansion of the
594 precuneus in human evolution. *Brain Struct. Funct.* **222**, 1053–1060 (2017).

595 61. Kochiyama, T. *et al.* Reconstructing the Neanderthal brain using computational
596 anatomy. *Sci. Rep.* **8**, 6296 (2018).

62. Neubauer, S., Hublin, J.-J. & Gunz, P. The evolution of modern human brain shape. *Sci Adv* **4**, eaao5961 (2018).

63. Pereira-Pedro, A. S., Bruner, E., Gunz, P. & Neubauer, S. A morphometric comparison of the parietal lobe in modern humans and Neanderthals. *J. Hum. Evol.* **142**, 102770 (2020).

64. Brady, W. J., Crockett, M. J. & Van Bavel, J. J. The MAD Model of Moral Contagion: The Role of Motivation, Attention, and Design in the Spread of Moralized Content Online. *Perspect. Psychol. Sci.* **15**, 978–1010 (2020).