Shifting evaluative construal:

Common and distinct neural components of moral, pragmatic, and hedonic evaluations

Clara Pretus
Department of Psychobiology and Methodology of Health Sciences
Universitat Autònoma de Barcelona
ORCID: 0000-0003-2172-1184
clara.pretus@uab.cat

Jillian K. Swencionis
Center for Policing Equity
Yale University
ORCID: 0000-0001-5372-9186

Yifei Pei
Department of Psychology and Center for Neural Science
New York University
ORCID: 0009-0005-9956-2226

Luis Marcos-Vidal
Department of Psychiatry and Forensic Medicine
Universitat Autònoma de Barcelona
ORCID: 0000-0003-2135-1478

Ingrid J. Haas
Department of Political Science and Center for Brain, Biology, and Behavior
University of Nebraska-Lincoln
ORCID: 0000-0002-7316-6605
ihaas2@unl.edu

William A. Cunningham
Department of Psychology
University of Toronto
ORCID: 0000-0002-7063-5859

Dominic J. Packer
Department of Psychology
Lehigh University
ORCID: 0000-0001-8420-6548

Jay J. Van Bavel
Department of Psychology and Center for Neural Science
New York University
ORCID: 0000-0002-2520-0442
jay.vanbavel@nyu.edu

Correspondence should be directed to Clara Pretus & Jay Van Bavel, Department of Psychology, New York University, New York, New York, 10003 (phone: 212-992-9627)

Author note

The authors declare no conflict of interest. The behavioral data, analysis code, and research materials are available in Open Science Framework. The neuroimaging data will be shared with other researchers upon request. This study's design and its analysis were not pre-registered. A preprint of this manuscript has been made available in PsyArXiv at 10.31234/osf.io/esby8.

Abstract (225/250)

People generate evaluations of different attitude objects based on their goals and aspects of the social context. Prior research suggests that people can shift between at least three types of evaluations to judge whether something is good or bad: *pragmatic* (how costly or beneficial it is), *moral* (whether it's aligned with moral norms), and *hedonic* (whether it feels good; Van Bavel et al., 2012). The current research examined the neurocognitive computations underlying these types of evaluations to understand how people construct affective judgments. Specifically, we examined whether different types of evaluations stem from a common neural evaluation system that incorporates different information in response to changing evaluation goals (moral, pragmatic, or hedonic), or distinct evaluation systems with different neurofunctional architectures. We found support for a hybrid evaluation system in which people rely on a set of brain regions to construct all three forms of evaluation but recruit additional distinct regions for each type of evaluation. The three types of evaluations all relied on common neural activity in affective structures such as the amygdala, the insula, and the hippocampus. However, moral evaluations involved greater neural activation in the orbitofrontal and cingulate cortex compared to pragmatic evaluations, and temporoparietal regions compared to hedonic evaluations. These results suggest that people use a hybrid system that includes common evaluation components as well as distinct ones to generate moral judgments.

Keywords (5 max.)

Evaluation; Moral evaluation; fMRI; Framing effects; Evaluative construal

Shifting evaluative construal: Common and distinct neural components of moral, pragmatic, and hedonic evaluations

"There is nothing either good or bad, but thinking makes it so."

- William Shakespeare (Hamlet)

From slime molds to Hamlet, determining whether something is good or bad is a critical faculty for any living organism. In humans, the evaluation of a particular entity with some degree of favor or disfavor is known as an attitude (Allport, 1935; Eagly & Chaiken, 2007). Attitudes enable people to determine, often quickly and effortlessly, how to act, which people to approach or avoid, and which products to buy (Duckworth et al., 2002; Maio & Olson, 2000). For the past century, psychologists have measured these evaluative preferences across a wide range of issues–making attitudes one of the most indispensable concepts in psychology (Allport, 1935; Thurstone, 1928). The most basic conceptualization of an attitude is a link between an object and positive and negative valence (Fazio, 2007). These simple associations drive automatic evaluations of a wide range of stimuli, including objects, actions, individuals, and social groups (Albarracin et al., 2014; Wood, 2003). However, evaluations can stem from a wide variety of considerations, appraisals, and cognitions beyond these immediate associations (Fazio, 2007; Schwarz, 2007). The current paper examines the online construction of evaluations, by examining the neural computations of different types of evaluations.

Although evaluations consist of valence judgments–*is something good or bad*–there is a longstanding debate about how people generate these evaluations for different attitude objects. For instance, attitudes appear to be made up of cognitive, affective, and behavioral components–with affective components often the most powerful (Abelson et al., 1982). In particular, attitudes that are moralized (i.e., grounded in a sense of right vs. wrong) appear to be very powerful (Skitka et al., 2021). These "moral mandates" are perceived as universally and objectively true, which can lead to greater intolerance of people who hold different attitudes, immunity to social influence, and increased political engagement. For instance, feelings of hatred toward people and concepts are more likely to be moralized than mere feelings of dislike (Pretus et al., 2023). In fact, merely labeling an attitude as moral increases its strength (Luttrell et al., 2016). As such, the process of moralizing issues can make evaluations more extreme.

Constructivist models of attitudes and appraisal theories explain how these affective evaluations emerge from multiple sources (Barrett, 2017; Cunningham et al., 2013; Schwarz & Bohner, 2001). According to these approaches, people generate evaluations of different attitude objects in the moment based on their goals and aspects of the social context (Scherer, 1997; Schwarz & Bohner, 2001). Differences in appraisal, in turn, lead to differences in emotional experiences. For instance, people may be pleased-displeased about outcomes relevant to their goals, approve-disapprove of actions related to social standards, and like-dislike objects based on their tastes (Clore & Huntsinger, 2009). These appraisals stem, in part, from the functions of different attitudes for the individual; for example, the functions of social adjustment, ego-defense, value expression, and knowledge function (Katz, 1960). From this perspective, attitudes reflect the given social needs of the perceiver rather than the stimulus itself. For instance, people like cake more when they are hungry than when they are not (Ferguson & Bargh, 2004). As such, understanding *what* perceivers hope to accomplish is critical to understanding *how* they construct their evaluations of attitude objects.

Prior research suggests that people can shift between at least three modes of evaluation: *pragmatic*, *hedonic*, and *moral* (Van Bavel et al., 2012). When people make moral judgments (rating whether actions are morally good or bad) their evaluations are faster, more extreme, and more strongly associated with universal prescriptions—the belief that absolutely nobody or everybody should engage in an action—than when they make pragmatic or hedonic evaluations of the same actions. For instance, people judge riding a bike in more extreme terms (seeing it as extremely good or bad) when they evaluate it through a moral lens. How do people arrive at such different evaluations of the same action from one moment to the next? We propose that each mode of evaluation elicits distinct cognitive operations (e.g., pragmatic likely involves more deliberate cost-benefit analyses, moral likely involves social relationships, and hedonics are about how people feel). Although there is good evidence that affective states influence valuation and choice (see Phelps et al., 2014), less is known about the mental computations underlying these specific modes of evaluation.

Overview

In the current research, we use neuroimaging to examine how different types of evaluations are constructed. Building on prior research, we used a social neuroscience approach to understand the neural architecture underlying different types of evaluation (Cacioppo et al., 2000; Cunningham & Zelazo, 2007). This approach provides insights into the different brain regions associated with distinct construals to provide insights into how people render different types of evaluations. In particular, we examined whether different types of evaluations are constructed based on a common neural evaluation system that incorporates different information in response to changing evaluative goals (moral, pragmatic, or hedonic), or distinct evaluation systems with distinct neural architectures. In neuroeconomics, the notion of common currency suggests that *different* types of actions are evaluated by the same neural system–located in the ventromedial prefrontal cortex (Cunningham et al., 2011; Levy & Glimcher, 2012; Ruff & Fehr, 2014). Here, we examine whether evaluations of the *same* actions might elicit distinct neural systems. Therefore, we kept the stimuli constant across conditions to isolate *how* people make different types of evaluations of the same actions (see also Cunningham et al., 2008). Unconfounding the attitude objects from the types of evaluation allowed us to see how people construct moral, pragmatic, and hedonic evaluations.

If different types of evaluation rely on different neurofunctional systems, then switching between different types of evaluations should recruit distinct brain regions associated with each evaluation type. For instance, moral evaluations could be associated with neural activation in brain areas associated with moral intuitions (e.g., the medial prefrontal cortex, Greene, 2007; Greene & Haidt, 2002), pragmatic evaluations could elicit activity in brain regions associated with deliberation and cost-benefit analysis (e.g., the dorsolateral prefrontal cortex, Weissman et al., 2008), and hedonistic evaluations could rely on neural activation in reward anticipation centers (the nucleus accumbens, Knutson & Greer, 2008).

Conversely, a common evaluation system would involve a common neuroarchitecture with little variation in neural activation across different types of evaluations. In addition to the role of the vmPFC and its role in computing a common currency (Cunningham et al., 2011; Levy & Glimcher, 2012), evaluative processes have been associated with neural activity in regions such as the insula, which is thought to integrate cognitive and affective processes, as well as the amygdala, a region linked to affective arousal (Berntson et al., 2011). In the current experiment, we examine whether different types of evaluations, even those traditionally dissociated from emotion (e.g., moral and pragmatic evaluations; Kohlberg, 1969; Piaget, 1932; Turiel, 1983), rely on common activation patterns in these affective systems, or instead, result from distinct evaluation systems.

Methods

This research was approved by the New York University Institutional Review Board according to the Declaration of Helsinki guidelines. The data collection took place in the Fall of 2010.

Transparency and openness

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study, and we follow JARS (Kazak, 2018). The behavioral data, analysis code, and research materials are available in Open Science Framework. The neuroimaging data will be shared with other researchers upon request. Data were analyzed using R, version 4.2.1 (R Core Team, 2022) and the package ggpubr, version 0.6.0 (Kassambara, 2023). This study's design and its analysis were not pre-registered.

Participants

A power analysis based on 2000 simulations and an alpha = 0.05 showed that with 20 participants and 252 observations per participant (84 trials per condition split into 7 runs), we would be able to detect small effects (e.g., a slope of 0.1) for the effect of evaluation type (moral/pragmatic/hedonic) with a statistical power of 100% (95% CI [99.82, 100]). The large number of trials in each condition allowed us to obtain a very good statistical power to identify behavioral differences between conditions and within participants.

Twenty-three participants (10 males, 13 females; mean age = 22.4 years, SD = 5.3) were recruited from the New York University community and paid $40 for completing the study.

Participants were right-handed, reported no abnormal neurological history, and had normal or corrected-to-normal vision. Demographic data on ethnicity, immigration history, and socioeconomic status were not collected.

Procedure

Practice trials. After giving written consent to participate in the experiment, participants completed a set of 45 practice trials on a laptop. Instructions for the practice task informed participants that they would evaluate a number of different actions in three different ways, and were given the following descriptions: *"Pragmatic evaluations focus on the costs and benefits you may experience if you do something. Moral evaluations focus on whether you ought to do something because it is right or wrong. Hedonic evaluations focus on whether an action would feel pleasant or unpleasant."* Participants were asked to consider the example of "getting a flu shot," which could be pragmatically good because it reduces the probability of illness, morally good because it reduces the spreading of illness to others, but hedonically bad because it is unpleasant. Then, participants began a set of practice trials in which they rated different actions in either moral, pragmatic, or hedonic terms on a 5-point scale, with 1 = *very bad* and 5 = *very good*[1] (Example: "How pragmatically bad/good would it be for you to ride a bike?", see Fig. 1). Before each block of trials (4 trials per block), participants were instructed, for example, "The next series of questions will focus on HEDONIC judgments".

Evaluation task during fMRI. After completing the practice trials, each participant was brought into the scan center and instructed to complete the same evaluation task while inside the

[1] After testing different endpoints in prior research (e.g., good/bad vs. right/wrong) we decided to use the same endpoints (very good/very bad) across the different evaluation conditions (moral/pragmatic/hedonic) to reduce potential differences associated with this wording (see Van Bavel et al., 2012).

scanner using a button box placed in the right hand to rate actions on a 5-point scale. During 7 functional runs in a blocked design, each participant completed ratings of 84 actions for each type of evaluation (moral, pragmatic, and hedonic) for a total of 252 trials. Each action was rated 3 times across the three types of evaluations. In each functional run, participants completed 3 blocks of each type of evaluation, in addition to 3 blocks in which they were instructed to rest, in randomized order. Thus, each functional run included 12 blocks split evenly between moral, pragmatic, hedonic, and rest. Each evaluation block contained an instruction screen (4 s) and a fixation cross (1 s) followed by either 4 moral, pragmatic, or hedonic trials in a row (each 5 s). Each trial was separated by a fixation cross (1s). The resting blocks involved a 30 s period of rest.

Fig. 1. fMRI task conditions. Participants were asked to evaluate how good or bad a series of 84 actions were from a moral, pragmatic, and hedonistic perspective.

Individual differences. After the scan session, participants completed individual differences scales on a laptop. These scales included a modified version of the Moral Mandates questionnaire (Mullen & Skitka, 2006) for each action rated in the scanner (M = 3.60, SD =0.43; “To what extent are your feelings about being punctual deeply connected to your beliefs about *right* and *wrong*?” with a 5-point Likert scale from 1 = *Not at all* to 5 = *Very much*); and a Moral Foundations scale (“When you decide whether something is right or wrong, to what extent are the following considerations relevant to your thinking?" with a 6-point Likert scale from 0 = *Not at all relevant* to 5 = *Very relevant*) (Haidt & Joseph, 2004) including subscales for authority (M = 3.50, SD = 0.75, α = .64, e.g., "Whether or not someone showed respect for authority"), loyalty items (M = 3.34, SD = 0.68, α = .56, "Whether or not someone showed a lack of loyalty"), purity (M = 3.60, SD = 0.43, α = .76, "Whether or not someone did something disgusting"), harm (M = 4.56, SD = 0.56, α = .44, "Whether or not someone suffered emotionally"), and fairness (M = 4.41, SD = 0.84, α = .84, "Whether or not someone acted unfairly").

MRI acquisition

Images were acquired in a Siemens 3T scanner at the NYU Center for Brain Imaging. T1-weighted images were obtained using a fast spoiled gradient-echo (FSPGR) sequence (176 slices, slice thickness: 1.00 mm, echo time: 3.93 msec, repetition time: 2500 msec, matrix size: 256 × 256, FA: 8). Functional scanning was prescribed +20 degrees parallel to the anterior commissure–posterior commissure (AC/PC) line, and nearly isotropic functional images were acquired in an interleaved manner from inferior to superior using a multi-echo pulse sequence (34 axial slices, 3 mm thick, 0.5mm skip, echo time = 15 msec, repetition time = 2000 msec, in-plane resolution = 3 x 3 x 3 mm, matrix size = 80 x 64, field of view = 192 x 80 mm). We obtained 176 volumes per run in 7 runs.

Data were preprocessed in SPM12 (Wellcome Department of Cognitive Neurology, London, United Kingdom) in MATLAB using a standard preprocessing pipeline in CONN toolbox (Nieto-Castanon & Whitfield-Gabrieli, 2021) that included functional realignment and unwarping, outlier identification, direct segmentation, normalization to MNI standard space, and functional smoothing using a Gaussian kernel of 8mm FWHM to reduce noise (see Nieto-Castanon, 2020 for more details). No slice-timing correction was performed. One participant was excluded due to incomplete data acquisition and one participant was excluded due to excessive head movement.

Behavioral analysis

To investigate the effect of evaluation type (moral/pragmatic/hedonic) on participants' judgments (good/bad) we conducted a series of mixed effects models with random intercepts for participants using REML (afex package in R, Singmann et al., 2016). We also assessed the effect of evaluation type on response latency and the extremity of participants' evaluations. For that, we computed an extremity score by converting the neutral midpoint of the 5-point response scale to 1 and the bipolar endpoints of the response scale to 3. To assess differences in the effect of type of evaluation on extremity scores as a function of individual differences in moral mandates and moral foundations we added an interaction term for moral mandates' scores and evaluation type.

Neuroimaging analysis

Image processing was conducted using SPM12 (Wellcome Trust, University College London, UK). The first-level general linear model (GLM) included one regressor with the hrf-convolved onset times for each evaluation type, including one moral, one pragmatic, and one hedonic evaluation regressor. The duration of these events was 5 seconds (trial duration). Each of the three task condition regressors was parametrically modulated by the participant's response to each trial on a 5-point Likert scale. Because responses to moral evaluations were faster than responses to pragmatic and hedonic evaluations, an additional parametric regressor for response latency was included for each type of event. Missing trials (no response throughout the trial, which was observed in 32 trials across all participants) were not included in the GLM and, thus, were processed as part of the implicit baseline. We also included a regressor for the instructions screen (shown for 4 seconds at the beginning of each of the 12 blocks per run), and a button press regressor to control for the confounding effects of motor preparation and finger movement. To control for head movement, we included scrubbing regressors that accounted for outlier volumes as required by each participant and six non-convolved head movement regressors. Outliers included volumes with a framewise displacement larger than 0.9 mm or global BOLD signal changes above 0.5 SD.

Group-level effects of the (a) moral vs. pragmatic evaluation, (b) moral vs. hedonic evaluation, and (c) pragmatic vs. hedonic evaluation were computed using a one-sample t-test on the individual-level contrasts of the onset regressors. Post hoc analyses included parameter estimate extraction in regions identified in the moral vs. pragmatic and moral vs. hedonic contrast, including right Middle Orbital Gyrus, right Middle Cingulate, and bilateral Supramarginal Gyrus activity, using Marsbar (Brett et al., 2002) to graphically represent the amount of neural activation in each condition. All reported results stem from whole-brain analyses and were corrected for multiple comparisons using family-wise error (FWE) correction at cluster-level, single voxel $P < 0.001$.

To examine common activity patterns between different types of evaluations (moral, pragmatic, hedonic), we generated binary masks of each type of evaluation, using the onset regressors, compared to baseline thresholded at single voxel $P < 0.05$ FWE-corrected. We then conducted a conjunction analysis using the binarized masks and obtained the overlapping brain regions that were active across the three types of evaluations (*versus* baseline).

Results

Behavioral results

Ratings across types of evaluations. Ratings for each action item and condition are displayed in Table S1. On average, ratings during the fMRI task did not differ between conditions (Moral vs. pragmatic: $B = -0.08$, 95% CI [-0.46, 0.29], $t(102) = -0.43$, $p = .67$; Moral vs. hedonic: $B = -0.08$, 95% CI [-0.54, 0.38], $t(102) = -0.34$, $p = 0.74$) and ratings of the same action items were highly correlated across conditions (Moral vs. pragmatic: $r = 0.84$, $p < 0.001$; moral vs. hedonic: $r = 0.84$, $p < 0.001$; pragmatic vs. hedonic: $r = 0.87$, $p < 0.001$). Thus, moral, pragmatic, and hedonic evaluations of the same actions yielded similar ratings across participants.

Extremity. Based on our prior findings (Van Bavel et al., 2012), we predicted that moral evaluations would be more extreme than pragmatic or hedonic evaluations. The mixed-effects model revealed a large effect of evaluation type on extremity scores (Moral vs. pragmatic: $B = 0.25$, 95% CI [0.17, 0.34], $t(40) = 10.83$, $p < .001$, Cohen's $d = 0.98$, 95% CI [0.32, 1.64]; Moral vs. hedonic: $B = 0.20$, 95% CI [0.11, 0.28], $t(40) = 8.63$, $p < .001$, $d = 0.66$, 95% CI [0.02, 1.30]), confirming this hypothesis (see Fig 2A). Pairwise comparisons revealed that mean moral extremity ($M = 2.40$, $SD = .72$) was significantly greater than pragmatic ($M = 2.15$, $SD = .73$) or hedonic ($M = 2.21$, $SD = .74$) extremity ($p < .001$ for both comparisons), but that pragmatic and hedonic extremity did not significantly differ ($p = .46$).

It is worth noting that we found no difference in the valence of actions (expected *good* or *bad*) on participants' extremity ratings across conditions ($B = 0.08$, 95% CI [-0.06, 0.23], $t(60) = 1.10$, $p = .277$). Overall, these results largely replicated prior work on moral construal (Van Bavel et al., 2012).

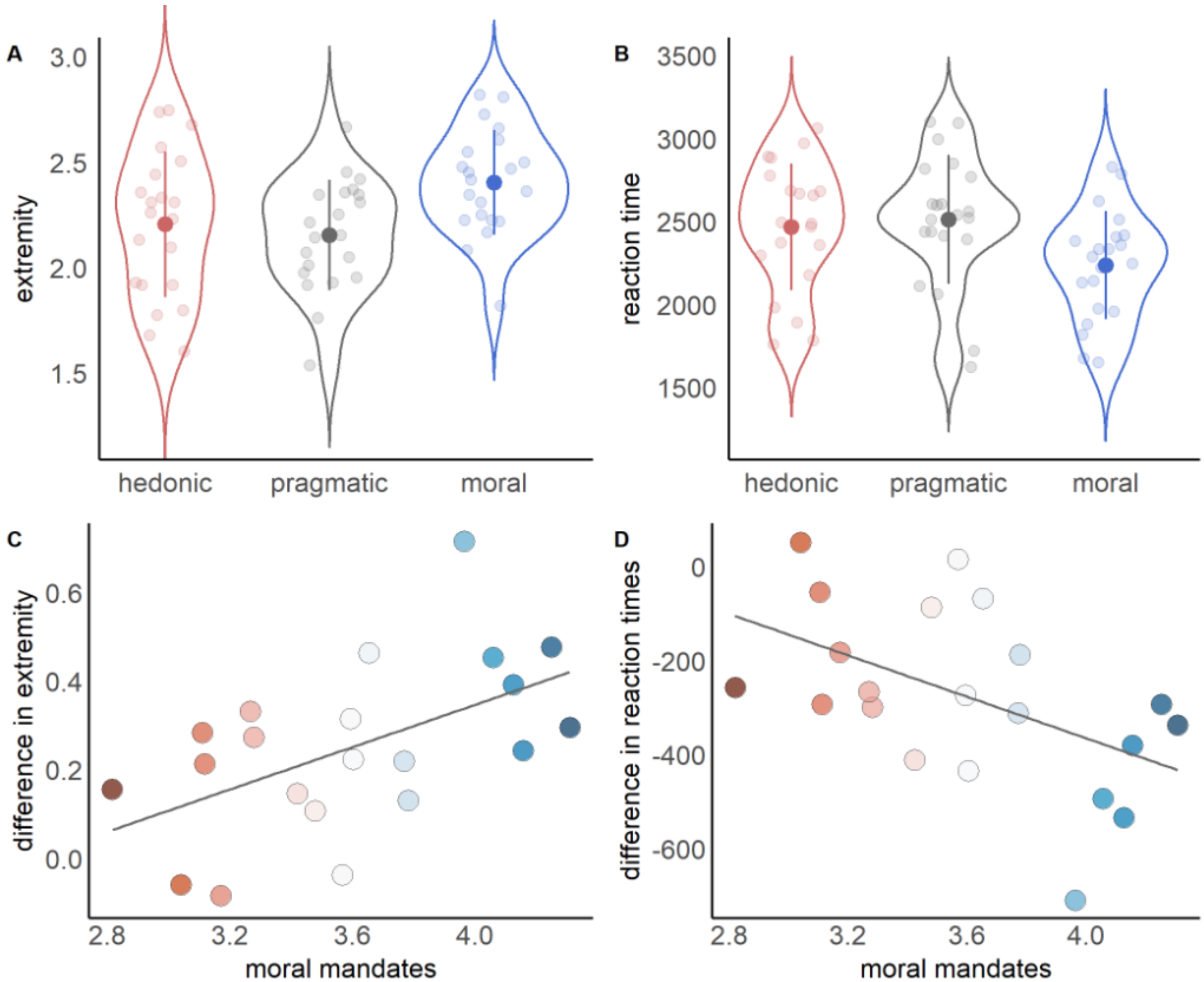

Fig 2. Behavioral results. (A) Extremity scores were higher for moral evaluation trials compared to pragmatic and hedonic trials during the fMRI task; **(B)** Response latency was faster for moral evaluation trials compared to pragmatic and hedonic trials during the fMRI task; **(C)** Participants who scored higher on moral mandates exhibited higher extremity scores in moral compared to pragmatic evaluation trials; and **(D)** Participants who scored higher on moral mandates exhibited shorter reaction time in moral compared to pragmatic evaluation trials.

Response Latency. We also predicted that moral evaluations would be made more quickly than pragmatic or hedonic evaluations. The mixed-effects model on response latency also revealed a large effect of condition (Moral vs. pragmatic: B = 275.74, 95% CI [200.44, 351.04], t(40) = 7.17, p < .001, d = 0.77, 95% CI [0.12, 1.41]; Moral vs. hedonic: B = 230.70, 95% CI [155.39, 305.99], t(40) = 6.00, p < .001, d = 0.65, 95% CI [0.01, 1.29]), confirming our hypothesis about response latency (see Fig 2B). Pairwise comparisons revealed that the mean reaction time for moral evaluations (M = 2236.03 ms, SD = 841.36) was significantly faster than pragmatic (M = 2511.75 ms, SD = 897.35) or hedonic (M = 2467.53 ms, SD = 930.59) response latency (*p* <.001 for both comparisons), but that pragmatic and hedonic response latency did not significantly differ (*p* = .48). These results were robust even after log-transforming response latency (Moral vs. pragmatic: B = 0.11, 95% CI [0.08, 0.14], t(40) = 7.20, p < .001; Moral vs. hedonic: B = 0.10, 95% CI [0.07, 0.12], t(40) = 6.10, p < .001). These results replicated prior work on moral construal (Van Bavel et al., 2012).

Moderation effects. To examine how people with different moral profiles responded to moral and non-moral evaluations, we examined whether individual differences in moral mandates and moral foundations influenced how quickly and extremely people responded to moral compared to non-moral evaluations in an exploratory analysis. We found that moral mandates (but not moral foundations) moderated both extremity scores (B = 0.24, 95% CI [0.06, 0.42], t(38) = 2.52, p = .016) and response latency (B = 220.41, 95% CI [57.03, 383.79], t(38) = 2.57, p = .014) in the moral *versus* pragmatic conditions. Specifically, participants with the highest moral mandates' scores (Mandates = 5) exhibited greater extremity scores (M_{diff} = 0.59, 95% CI [0.25, 0.92], t(38) = 4.22, p < .001, see Fig 2C) and faster response latency in moral compared to pragmatic evaluation trials (M_{diff} = 585.0, 95% CI [278.92, 891.20], t(38) = 4.66, p < .001, see Fig 2D) than participants with the lowest moral mandates' scores (Mandates = 1) (extremity: p = 0.31, and response latency: p = 0.40). Thus, participants who reported that the actions were more connected to their moral beliefs also responded more quickly and more extremely when making moral compared to pragmatic evaluations of those actions.

Neuroimaging results

Common neural activity. To test the hypothesis of a common neural evaluation system, we obtained the overlapping neural response of each type of evaluation by means of a conjunction analysis. Because the conjunction analysis was conducted using binarized masks ($P < 0.05$ FWE at a voxel level), we only report cluster extent (k). The common neural activation across moral, pragmatic, and hedonic evaluations was largely left-lateralized (see Fig. 2A) and was found in subcortical affective areas including the left amygdala ($k = 77$), the hippocampus ($k = 65$), and the thalamus ($k = 91$), and affective cortical regions such as the left insula ($k = 117$) and the Cerebellum and fusiform gyrus ($k = 4116$). We also found common activation in frontoparietal regions including the left superior medial gyrus ($k = 262$), the left inferior frontal gyrus (pars opercularis/triangularis, $k = 1213$; and pars orbitalis, $k = 216$), and the left inferior parietal lobule ($k = 814$). Thus, evidence of a common neural substrate for moral, pragmatic, and hedonic evaluations was found in both affective and frontoparietal regions typically involved in executive control.

Differential neural activity. To test the distinct evaluation systems hypothesis, we focused on the neural activity differences between types of evaluations. Comparing neural activity in *moral* compared to *pragmatic* evaluation trials across the whole brain revealed two clusters of activation in brain regions associated, among other functions, with integrating affective value into action planning (Ruff & Fehr, 2014; Rolls, 2019), including the right Middle Orbital Gyrus extending into the right Superior Orbital Gyrus and the left Anterior Cingulate ($k = 497$, $T > 4.39$, $p < .001$

FWEc), and the right Middle Cingulate (k = 182, $T = 5.75$, $p = .031$ FWEc, see Fig. 2B and Table 1a). Conversely, the contrast between *moral* and *hedonic* evaluation trials yielded neural activation in three different clusters in brain regions previously associated with social cognition (Saxe & Wexler, 2005; Hoffman et al., 2016), including the right Supramarginal Gyrus (k = 771, $T > 4.16$, $p < .001$ FWEc), the left Supramarginal Gyrus (k = 422, $T = 6.04$, $p = .001$ FWEc), and the Middle Temporal Gyrus (k = 200, $T > 3.87$, $p = .027$ FWEc, see Fig. 2C and Table 1b). The contrast between *pragmatic* and *hedonic* evaluations did not yield any significant results ($p > 0.98$ FWEc). Thus, we identified different neural activity patterns in response to moral and non-moral evaluations of the same actions but were unable to detect differences between pragmatic and hedonic evaluations using the present experimental design.

As an exploratory analysis, we examined neural activity differences between different types of actions by classifying them based on whether they appealed to any of the five moral foundations (e.g., “listen to parents” and authority). However, this analysis yielded no significant results (see Supplementary results).

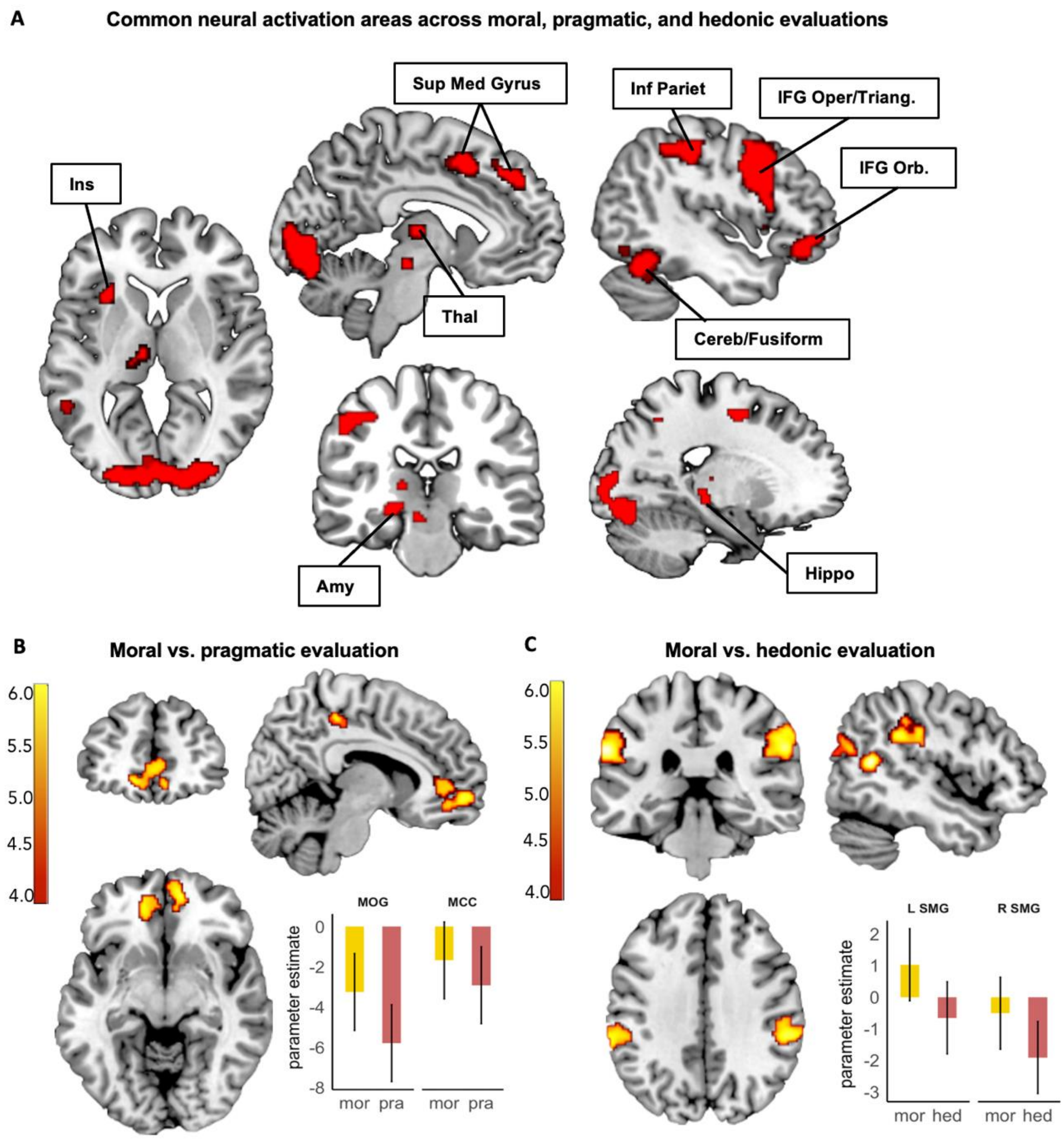

Fig 2. Neuroimaging results. (A) Common neural activation areas across different types of evaluations, including affective areas such as the left insula (Ins), the amygdala (Amy), the hippocampus (Hippo), the thalamus (Thal), and the cerebellum/fusiform gyrus (Cereb/Fusiform), and frontoparietal regions such as the superior medial gyrus (Sup Med Gyrus), inferior parietal lobule (Inf Pariet), inferior frontal gyrus (IFG) pars opercularis (Oper.), pars triangularis (Triang.), and pars orbitalis (Orb.). **(B)** Moral evaluations ("mor") compared to pragmatic evaluations ("pra") of the same actions yielded greater neural activation in the right Middle Orbital Gyrus (MOG) and

the right Middle Cingulate Cortex (MCC); **(C)** Moral evaluations ("mor") compared to hedonic evaluations ("hed") predicted higher neural activation in the bilateral Superior Marginal Gyrus (SMG). Parameter estimates in **(B)** and **(C)** have been calculated using a functional mask of the right Middle Orbital Gyrus (MOG) and the right Middle Cingulate Cortex (MCC) activity, and the bilateral Superior Marginal Gyrus (SMG), respectively, extracted from the displayed contrast.

Table 1. Neuroimaging results table. Differential neural activity detected in response to each type of evaluation and as a function of purity and authority scores of the Moral Foundations Questionnaire.

			MNI coordinates			
Regional label	**k**	**t(20)**	**x**	**y**	**z**	**P-FWEc**
(a) Moral vs. pragmatic evaluation						
R Middle Cingulate	182	5.75	16	-24	44	.031
R Middle Orbital Gyrus	497	4.81	10	54	-6	< .001
L Superior Orbital Gyrus	497	4.44	-14	44	-8	< .001
R Anterior Cingulate	497	4.39	2	36	6	< .001
(b) Moral vs. hedonic evaluation						
R Middle Temporal Gyrus	200	6.08	46	-54	12	.027
R Middle Temporal Gyrus	200	3.87	48	-72	28	.027
L SupraMarginal Gyrus	422	6.04	-62	-30	30	.001
R SupraMarginal Gyrus	771	5.65	58	-32	32	< .001
R Rolandic Operculum	771	4.16	38	-20	26	< .001

Discussion

We investigated whether different types of evaluations (moral, pragmatic, and hedonic) stem from a common evaluation system that incorporates different information in response to changing evaluation goals, or distinct evaluation systems with distinct neurofunctional architectures. We found support for a hybrid evaluation system: moral, pragmatic, and hedonic evaluations relied on common neural substrates in affective regions (e.g., amygdala and insula) as well as frontoparietal regions typically involved in executive control (e.g., inferior frontal gyrus), but moral evaluations additionally elicited distinct neural activations compared to non-moral evaluations. Compared to pragmatic evaluations, moral evaluations were associated with greater activity in regions responsible for integrating social reward and affective value into subsequent action planning, among other functions, such as the orbitofrontal and middle cingulate cortex. Compared to hedonic evaluations, moral evaluations elicited greater activity in regions previously associated with social cognition such as the bilateral temporoparietal junction. Behaviorally, moral evaluations were also different from non-moral evaluations in at least two dimensions: they elicited faster and more extreme responses than pragmatic and hedonic evaluations, especially among participants who reported that the evaluated actions were connected to their moral beliefs. Thus, both our behavioral and neural findings are compatible with a hybrid evaluation system with common neurofunctional correlates (e.g., in affective regions) and additional specific neural activation (e.g., in the orbitofrontal cortex) for at least one type of evaluation: moral evaluations.

Our findings support the idea of a hybrid evaluation system with several common neural correlates across moral, pragmatic, and hedonic evaluations. This common neurofunctional architecture is markedly left-lateralized and includes brain regions previously associated with affective processes (i.e., the amygdala and insula, Berntson et al., 2011) and action-selection prediction (i.e., the hippocampus, Johnson et al., 2007) during evaluation. Some frontoparietal activity was also identified across types of evaluation. However, it is difficult to disentangle how much of this neural activity is related to the overall evaluation process or is specific to the employed neuroimaging task. For instance, using a Neurosynth association analysis (Yarkoni et al., 2011), we found that the coordinates of the neural activation obtained in the inferior frontal gyrus and the superior medial gyrus have been reported in relation to language processing, while the coordinates of the inferior parietal lobule have been linked to working memory. Thus, while we cannot provide a comprehensive picture of the common neural correlates across different types of evaluations, we found evidence for at least several common components in affective regions, including the amygdala, the insula, and the hippocampus.

In addition to the common components, we found specific neural activation for moral compared to non-moral evaluations. As compared to pragmatic evaluations, moral evaluations were associated with activity in the orbitofrontal and middle cingulate regions of the brain. The orbitofrontal cortex has been previously associated with prepotent affective responses in personal moral dilemmas (Greene et al., 2001), subjective value computations during decision-making (Ruff & Fehr, 2014), representation of rewards during social decision-making (Nakamura et al., 2008; O'Doherty et al., 2001), and pro-social moral sentiments (Moll & de Oliveira-Souza, 2007). Meanwhile, the middle cingulate cortex has been involved in action-outcome learning (Rolls, 2019). This region receives reward and affective value inputs from the orbitofrontal cortex via the anterior cingulate (Rolls & Grabenhorst, 2008), as well as visuospatial and reward-related episodic memory inputs from the parietal cortex and the hippocampus via de posterior cingulate (Rolls & Wirth, 2018). The middle cingulate is thought to integrate this emotional information to generate motor behavior via outputs to motor and premotor regions involved in goal-directed behavior

(Oane et al., 2020; Shackman et al., 2011). Thus, moral evaluations seem to be supported by neural networks responsible for detecting and integrating social reward and affective value into subsequent action planning, while pragmatic evaluations may be less reliant on social reward computations.

As compared to hedonic evaluations, moral evaluations elicited greater activity in the bilateral supramarginal gyrus and the right middle temporal gyrus. The left supramarginal gyrus has been associated with rapid action reprogramming (Hartwigsen et al., 2012) and planning and observation of tool-directed actions (Potok et al., 2019). Meanwhile, the right supramarginal gyrus has been involved in emotion recognition (Wada et al., 2021), overcoming egocentric bias in social judgments (Silani et al., 2013), and self-other distinction in empathy tasks (Hoffman et al., 2016). Together with the right middle temporal gyrus, the right supramarginal gyrus forms the temporoparietal junction, which has been associated with attribution of mental states (Saxe & Wexler, 2005), relying on others' mental states in moral judgments (Young et al., 2010) and enabling a sense of agency based on self-other distinctions (Quesque & Brass, 2019). Thus, our results suggest that moral evaluations may rely on neural activity patterns associated with mental state attribution and self-other distinction during social and moral judgments, while hedonic evaluations may be the result of more self-centered computations.

Our findings add nuance to existing literature on the neural underpinnings of moral judgments. For instance, moral tasks have been associated with common neural activation in the orbitofrontal cortex, the temporoparietal junction, and the amygdala in a meta-analysis based on 123 datasets (Eres et al., 2018). Our findings partially align with these results. On one hand, brain regions previously associated with affective appraisal during moral judgments such as the amygdala (Eres et al., 2018; Shenhav & Greene, 2014a) and the insula (Hutcherson et al., 2015) activated across all types of evaluations in our study, not only moral ones. On the other hand, brain regions previously associated with utilitarian appraisal such as the temproparietal junction (Eres et al., 2018; Hutcherson et al., 2015) and overall integration of moral value such as the orbitofrontal cortex (Hutcherson et al., 2015; Shenhav & Greene, 2014b) were specific to moral as compared to hedonic and pragmatic evaluations, respectively. This contrasts with studies suggesting that the orbitofrontal cortex serves a common currency function across different types of evaluations (Ruff & Fehr, 2014). Overall, our findings suggest that part of the affect-related regions associated with moral judgments are shared with other types of evaluations (e.g., subcortical regions such as the amygdala), while other affect-related regions are specific to moral evaluations compared to pragmatic ones (e.g., higher-order cortical regions such as the orbitofrontal cortex).

Our findings have implications for the instrumental account of emotion regulation, the idea that people not only regulate their emotions to feel good in the short-term, but also upregulate unpleasant emotions that promote long-term goals (Tamir, 2009). For instance, previous studies have found that people choose to engage in anger-inducing activities when they anticipate being in a confrontational context (Tamir et al., 2008). Thus, emotion regulation can serve both hedonic and instrumental goals. We find that hedonic and pragmatic evaluations elicit common activation in brain regions associated with affective appraisal, including the amygdala and the insula. Hence, our findings suggest that affect-related neural processes play a key role in evaluations related to utility, in line with the view that instrumental motives can induce emotional adjustments.

Constraints on generality

Our results are bound by some limitations. Firstly, we conducted a lab experiment that aimed to identify the commonalities and differences between moral, pragmatic, and hedonic evaluations by asking people to switch between the different types of evaluations. While people may direct attention to different stimulus features in ecological contexts, future research should address how people spontaneously evaluate the same behaviors in moral, pragmatic, and hedonic terms, in response to their personal goals and situational context. Secondly, while the employed task allowed us to identify differences between types of evaluation in a controlled way (keeping the stimuli constant), it did not allow us to clearly distinguish commonalities across types of evaluations from task-specific effects. Future studies should include a specific control condition to isolate effects of interest common across different types of evaluations from task-specific effects related to stimulus presentation and the response setting. Thirdly, we used non-jittered 1-second inter-trial intervals (ITIs) in our fMRI task, which could have reduced the efficacy of the design (Dale, 1999). A more efficient design with jittered ITIs could have yielded stronger results but would not have substantially altered the pattern of results. Finally, our sample comprises undergraduate students and university community members in the US, limiting the generalizability of the reported results to populations with a lower educational level, outside of the US, or different levels of need for cognition (Cacciopo et al., 1996) and emotional granularity (Barrett, 2006). Limitations related to the characteristics of our participants do not only concern the reported results, but also the procedure and materials employed. For instance, the stimuli employed were selected for the population tested and may not be directly transferable to other cultural contexts (e.g., “keep a lost IPod”). The labels “moral”, “pragmatic”, and “hedonic” may also be understood differently across cultures. Thus, future research should examine whether differences and commonalities in brain activity across different types of evaluations can be generalized across different cultural contexts in a meaningful way.

Conclusion

In sum, our experiment addresses whether moral, pragmatic, and hedonic evaluations rely on a common evaluation system or neurally distinct evaluation systems. We found evidence for a hybrid model, where part of the neural activity in subcortical and cortical affective regions such as the amygdala, the insula, and the hippocampus is shared across different evaluation modalities. We also found modality-specific neural activation for moral compared to non-moral evaluations, particularly in orbitofrontal, cingulate, and temporoparietal regions. Thus, moral and non-moral evaluations seem to rely on both shared and specific neural substrates.

Author contributions

J.V.B supervised, conceptualized, and conducted the investigation. C.P. curated the data, conducted the formal analysis, and prepared the original draft. Y.P. and L.M. curated the data and participated in the formal analysis. J.K.S. prepared, administered, and conducted the investigation. I.J.H., D.P., and W.C. contributed to conceptualization. All authors reviewed and edited the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Funding

This work was supported by the NYU Center for Brain Imaging and a European Union grant (grant no. 101070930). Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Innovation Council. Neither the European Union nor the granting authority can be held responsible for them.

Acknowledgments

We thank everyone who participated in this study.

References

Abelson, R. P., Kinder, D. R., Peters, M. D., & Fiske, S. T. (1982). Affective and semantic components in political person perception. *Journal of Personality and Social Psychology*, *42*(4), 619–630. https://doi.org/10.1037/0022-3514.42.4.619

Albarracin, D., Johnson, B. T., & Zanna, M. P. (2014). *The Handbook of Attitudes*. Psychology Press.

Allport, G. W. (1935). Attitudes. In *A Handbook of Social Psychology* (pp. 798–844). Clark University Press. https://psycnet.apa.org/record/1935-19907-010

Barrett, L. F. (2017). How emotions are made: The secret life of the brain. In *How emotions are made: The secret life of the brain.* (pp. xv, 425–xv, 425). Houghton Mifflin Harcourt.

Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and social psychology review*, *10*(1), 20-46.

Berntson, G. G., Norman, G. J., Bechara, A., Bruss, J., Tranel, D., & Cacioppo, J. T. (2011). The insula, the amygdala and evaluative processes. *Psychological Science*, *22*(1), 80. https://doi.org/10.1177/0956797610391097

Cacioppo, J. T., Berntson, G. G., Sheridan, J. F., & McClintock, M. K. (2000). Multilevel integrative analyses of human behavior: Social neuroscience and the complementing nature of social and biological approaches. In *Psychological Bulletin* (Vol. 126, Issue 6, pp. 829–843). American Psychological Association. https://doi.org/10.1037/0033-2909.126.6.829

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological bulletin*, *119*(2), 197.

Clore, G. L., & Huntsinger, J. R. (2009). How the Object of Affect Guides its Impact. *Emotion Review*, *1*(1), 39–54. https://doi.org/10.1177/1754073908097185

Cunningham, W. A., Dunfield, K. A., & Stillman, P. E. (2013). Emotional States from Affective Dynamics. *Emotion Review*, *5*(4), 344–355. https://doi.org/10.1177/1754073913489749

Cunningham, W. A., Johnsen, I. R., & Waggoner, A. S. (2011). Orbitofrontal cortex provides cross-modal valuation of self-generated stimuli. *Social Cognitive and Affective Neuroscience*, *6*(3), 286–293. https://doi.org/10.1093/SCAN/NSQ038

Cunningham, W. A., Van Bavel, J. J., & Johnsen, I. R. (2008). Affective Flexibility. *Psychological Science*, *19*(2), 152–160. https://doi.org/10.1111/J.1467-9280.2008.02061.X

Cunningham, W. A., & Zelazo, P. D. (2007). *Attitudes and evaluations : a social cognitive neuroscience perspective*. *11*(3). https://doi.org/10.1016/j.tics.2006.12.005

Dale A. M. (1999). Optimal experimental design for event-related fMRI. Human brain mapping, 8(2-3), 109–114.

Duckworth, K. L., Bargh, J. A., Garcia, M., & Chaiken, S. (2002). The automatic evaluation of novel stimuli. *Psychological Science*, *13*(6), 513–519. https://doi.org/10.1111/1467-9280.00490

Eagly, A. H., & Chaiken, S. (2007). The advantages of an inclusive definition of attitude. *Social Cognition*, *25*(5), 582–602. https://doi.org/10.1521/SOCO.2007.25.5.582

Eres, R., Louis, W. R., & Molenberghs, P. (2018). Common and distinct neural networks involved in fMRI studies investigating morality: an ALE meta-analysis. *Social Neuroscience*, *13*(4), 384–398. https://doi.org/10.1080/17470919.2017.1357657

Fazio, R. H. (2007). Attitudes as object-evaluation associations of varying strength. In *Social Cognition* (Vol. 25, Issue 5, pp. 603–637). Guilford Publications. https://doi.org/10.1521/soco.2007.25.5.603

Ferguson, M. J., & Bargh, J. A. (2004). Liking is for doing: the effects of goal pursuit on automatic evaluation. *Journal of Personality and Social Psychology*, *87*(5), 557–572. https://doi.org/10.1037/0022-3514.87.5.557

Greene, J. D. (2007). Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends in Cognitive Sciences*, *11*(8), 322–323. https://doi.org/10.1016/J.TICS.2007.06.004

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Joshua, D., Brian, R., Leigh, E., John, M., & Jonathan, D. (2001). An fMRI Investigation of Emotional Engagement in Moral Judgment. *Science*, *293*(5537), 2105–2109.

Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, *6*(12), 517–523. https://doi.org/10.1016/S1364-6613(02)02011-9

Haidt, J., & Joseph, C. (2004). Intuitive Ethics: How Innately Prepared Intuitions Generate Culturally Variable Virtues. *Daedalus*, *133*(4), 55–66. http://www.jstor.org/stable/20027945

Hartwigsen, G., Bestmann, S., Ward, N. S., Woerbel, S., Mastroeni, C., Granert, O., & Siebner, H. R. (2012). Left Dorsal Premotor Cortex and Supramarginal Gyrus Complement Each Other during Rapid Action

Reprogramming. *Journal of Neuroscience*, *32*(46), 16162–16171. https://doi.org/10.1523/JNEUROSCI.1010-12.2012

Hoffman, K. M., Trawalter, S., Axt, J. R., & Oliver, M. N. (2016). Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proceedings of the National Academy of Sciences*, *113*(16), 4296–4301. https://doi.org/10.1073/pnas.1516047113

Hutcherson, C. A., Montaser-Kouhsari, L., Woodward, J., & Rangel, A. (2015). Emotional and utilitarian appraisals of moral dilemmas are encoded in separate areas and integrated in ventromedial prefrontal cortex. *Journal of Neuroscience*, *35*(36), 12593–12605. https://doi.org/10.1523/JNEUROSCI.3402-14.2015

Johnson, A., van der Meer, M. A., & Redish, A. D. (2007). Integrating hippocampus and striatum in decision-making. *Current Opinion in Neurobiology*, *17*(6), 692. https://doi.org/10.1016/J.CONB.2008.01.003

Kassambara, A. (2023). *ggpubr: "ggplot2" Based Publication Ready Plots*. https://rpkgs.datanovia.com/ggpubr/

Katz, D. (1960). The funcional approach to the study of attitudes. *Public Opinion Quarterly*, *24*(2), 163–204. https://doi.org/10.1086/266945

Kazak, A. E. (2018). Editorial: Journal article reporting standards. In *American Psychologist* (Vol. 73, Issue 1, pp. 1–2). American Psychological Association. https://doi.org/10.1037/amp0000263

Knutson, B., & Greer, S. M. (2008). Anticipatory affect: neural correlates and consequences for choice. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1511), 3771–3786. https://doi.org/10.1098/RSTB.2008.0155

Kohlberg, L. (1969). Stage and Sequence: The Cognitive Developmental Approach to Socialization. In D. Goslin (Ed.), *Handbook of Socialization Theory and Research* (pp. 347–480). Rand McNally.

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology*, *22*(6), 1027–1038. https://doi.org/10.1016/j.conb.2012.06.001

Luttrell, A., Petty, R. E., Briñol, P., & Wagner, B. C. (2016). Making it moral: Merely labeling an attitude as moral increases its strength. *Journal of Experimental Social Psychology*, *65*, 82–93. https://doi.org/https://doi.org/10.1016/j.jesp.2016.04.003

Maio, G. R., & Olson, J. M. (2000). Why we evaluate: Functions of attitudes. In G. R. Maio & J. M. Olson (Eds.), *Why we evaluate: Functions of attitudes.* Lawrence Erlbaum Associates Publishers.

Moll, J., & de Oliveira-Souza, R. (2007). Moral judgments, emotions and the utilitarian brain. *Trends in Cognitive*

Sciences, *11*(8), 319–321. https://doi.org/10.1016/J.TICS.2007.06.001

Mullen, E., & Skitka, L. J. (2006). Exploring the psychological underpinnings of the moral mandate effect: Motivated reasoning, identification, or affect? *Journal of Personality and Social Psychology*, *90*, 629–643.

Nakamura, M., Nestor, P. G., Levitt, J. J., Cohen, A. S., Kawashima, T., Shenton, M. E., & McCarley, R. W. (2008). Orbitofrontal volume deficit in schizophrenia and thought disorder. *Brain*, *131*(1), 180–195. https://doi.org/10.1093/BRAIN/AWM265

Nieto-Castanon, A. (2020). FMRI minimal preprocessing pipeline. In Handbook of functional connectivity Magnetic Resonance Imaging methods in CONN (pp. 3–16). Hilbert Press. doi:10.56441/hilbertpress.2207.6599

Nieto-Castanon, A., & Whitfield-Gabrieli, S. (2021). CONN functional connectivity toolbox: RRID SCR_009550, release 21. *CONN Functional Connectivity Toolbox: RRID SCR_009550, Release 21*. https://doi.org/10.56441/HILBERTPRESS.2161.7292

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, *4*(1), 95–102. https://doi.org/10.1038/82959

Oane, I., Barborica, A., Chetan, F., Donos, C., Maliia, M. D., Arbune, A. A., Daneasa, A., Pistol, C., Nica, A. E., Bajenaru, O. A., & Mindruta, I. (2020). Cingulate cortex function and multi-modal connectivity mapped using intracranial stimulation. *NeuroImage*, *220*, 117059. https://doi.org/10.1016/J.NEUROIMAGE.2020.117059

Phelps, E. A., Lempert, K. M., & Sokol-Hessner, P. (2014). Emotion and Decision Making: Multiple Modulatory Neural Circuits. *Annual Review of Neuroscience*, *37*, 263–287. https://doi.org/10.1146/ANNUREV-NEURO-071013-014119

Piaget, J. (1932). *The Moral Judgment of the Child*. The Free Press.

Potok, W., Maskiewicz, A., Króliczak, G., & Marangon, M. (2019). The temporal involvement of the left supramarginal gyrus in planning functional grasps: A neuronavigated TMS study. *Cortex*, *111*, 16–34. https://doi.org/10.1016/J.CORTEX.2018.10.010

Pretus, C., Ray, J. L., Granot, Y., Cunningham, W. A., & Van Bavel, J. J. (2023). The psychology of hate: Moral concerns differentiate hate from dislike. *European Journal of Social Psychology*, *53*(2), 336–353. https://doi.org/10.1002/EJSP.2906

Quesque, F., & Brass, M. (2019). The Role of the Temporoparietal Junction in Self-Other Distinction. *Brain*

Topography, *32*(6), 943–955. https://doi.org/10.1007/S10548-019-00737-5/METRICS

R Core Team. (2022). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.r-project.org/.

Rolls, E. T. (2019). The cingulate cortex and limbic systems for emotion, action, and memory. *Brain Structure and Function*, *224*(9), 3001–3018. https://doi.org/10.1007/S00429-019-01945-2/FIGURES/4

Rolls, E. T., & Grabenhorst, F. (2008). The orbitofrontal cortex and beyond: From affect to decision-making. *Progress in Neurobiology*, *86*(3), 216–244. https://doi.org/10.1016/J.PNEUROBIO.2008.09.001

Rolls, E. T., & Wirth, S. (2018). Spatial representations in the primate hippocampus, and their functions in memory and navigation. *Progress in Neurobiology*, *171*, 90–113. https://doi.org/10.1016/J.PNEUROBIO.2018.09.004

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nat Rev Neurosci*, *15*(8), 549–562. http://dx.doi.org/10.1038/nrn3776%5Cn10.1038/nrn3776

Saxe, R., & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia*, *43*(10), 1391–1399. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2005.02.013

Scherer, K. R. (1997). The role of culture in emotion-antecedent appraisal. In *Journal of Personality and Social Psychology* (Vol. 73, Issue 5, pp. 902–922). American Psychological Association. https://doi.org/10.1037/0022-3514.73.5.902

Schwarz, N. (2007). Attitude Construction: Evaluation in Context. *Social Cognition*, *25*(5), 638–656.

Schwarz, N., & Bohner, G. (2001). The Construction of Attitudes. In A. Tesser & N. Schwarz (Eds.), *Blackwell Handbook of Social Psychology* (pp. 436–457). Blackwell.

Shackman, A. J., Salomons, T. V, Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews Neuroscience*, *12*(3), 154–167. https://doi.org/10.1038/nrn2994

Shenhav, A., & Greene, J. D. (2014a). Integrative moral judgment: dissociating the roles of the amygdala and ventromedial prefrontal cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *34*(13), 4741–4749. https://doi.org/10.1523/JNEUROSCI.3390-13.2014

Shenhav, A., & Greene, J. D. (2014b). Integrative Moral Judgment: Dissociating the Roles of the Amygdala and Ventromedial Prefrontal Cortex. *Journal of Neuroscience*, *34*(13), 4741–4749. https://doi.org/10.1523/JNEUROSCI.3390-13.2014

Silani, G., Lamm, C., Ruff, C. C., & Singer, T. (2013). Right Supramarginal Gyrus Is Crucial to Overcome Emotional Egocentricity Bias in Social Judgments. *Journal of Neuroscience*, *33*(39), 15466–15476. https://doi.org/10.1523/JNEUROSCI.1488-13.2013

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2016). *afex: Analysis of Factorial Experiments. R package version 0.16-1*. https://cran.r-project.org/package=afex

Skitka, L. J., Hanson, B. E., Morgan, G. S., & Wisneski, D. C. (2021). The Psychology of Moral Conviction. *Annual Reivew of Psychology*, *72*, 347–366. https://doi.org/10.1146/ANNUREV-PSYCH-063020-030612

Tetlock, P. E. (2002). Social functionalist frameworks for judgment and choice: Intuitive politicians, theologians, and prosecutors. *Psychological Review*, *109*(3), 451–471. https://doi.org/10.1037/0033-295X.109.3.451

Thurstone, L. L. (1928). Attitudes Can Be Measured. *American Journal of Sociology*, *33*, 529–554.

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge University Press.

Van Bavel, J. J., Packer, D. J., Haas, I. J., & Cunningham, W. A. (2012). The Importance of Moral Construal: Moral versus Non-Moral Construal Elicits Faster, More Extreme, Universal Evaluations of the Same Actions. *PLoS ONE*, *7*(11), e48693. https://doi.org/10.1371/journal.pone.0048693

Wada, S., Honma, M., Masaoka, Y., Yoshida, M., Koiwa, N., Sugiyama, H., Iizuka, N., Kubota, S., Kokudai, Y., Yoshikawa, A., Kamijo, S., Kamimura, S., Ida, M., Ono, K., Onda, H., & Izumizaki, M. (2021). Volume of the right supramarginal gyrus is associated with a maintenance of emotion recognition ability. *PLOS ONE*, *16*(7), e0254623. https://doi.org/10.1371/JOURNAL.PONE.0254623

Weissman, D. H., Perkins, A. S., & Woldorff, M. G. (2008). Cognitive control in social situations: A role for the dorsolateral prefrontal cortex. *NeuroImage*, *40*(2), 955–962. https://doi.org/10.1016/J.NEUROIMAGE.2007.12.021

Wood, W. (2003). Attitude Change: Persuasion and Social Influence. *Annual Review of Psychology*, *51*, 539–570. https://doi.org/10.1146/ANNUREV.PSYCH.51.1.539

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). NeuroSynth: a new platform for large-scale automated synthesis of human functional neuroimaging data. *Frontiers in Neuroinformatics*. http://neurosynth.org/

Young, L., Camprodon, J. A., Hauser, M., Pascual-Leone, A., & Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments.

Proceedings of the National Academy of Sciences of the United States of America, *107*(15), 6753–6758. https://doi.org/10.1073/PNAS.0914826107/SUPPL_FILE/PNAS.200914826SI.PDF