

Running head: WARNING BELL

Warning Bell: Liberals Implicitly Respond to Group Morality Before Rejecting it Explicitly

Jesse Graham

University of Southern California

Zoe Englander, James P. Morris, Carlee Beth Hawkins, Jonathan Haidt, & Brian A. Nosek

University of Virginia

May 31, 2012

Corresponding Author:

Jesse Graham

Department of Psychology

University of Southern California

3620 McClintock Ave.

Los Angeles, CA 90089

**Abstract**

Moral Foundations Theory (Haidt & Graham, 2007) explains the intractability of many political disagreements as the result of liberals and conservatives reacting to different patterns of intuitive moral concerns. Research using self-report measures has shown that liberals endorse the Care/harm and Fairness/cheating foundations but generally do not endorse the Loyalty/betrayal, Authority/subversion, and Sanctity/degradation foundations, while conservatives endorse all five types of moral concerns (Graham, Haidt, & Nosek, 2009). Research on ideology and implicit attitudes (Jost, Federico, & Napier, 2010; Jost, Nosek, & Gosling, 2008) suggests that liberals may have more of a discrepancy between implicit reactions and endorsed attitudes than do conservatives. The present studies test whether liberals have implicit reactions to moral stimuli that contradict their consciously endorsed attitudes. The studies take a multi-method approach in order to measure moral reactions using different foundation-related stimuli and procedures, including self-reported gut reactions (Study 1), evaluative priming (Study 2), the Affect Misattribution Procedure (Study 3), and EEG event-related potentials (Study 4). The studies provide convergent evidence of implicit group-focused moral concerns in liberals not captured by explicit measures.

Keywords: moral intuitions, political ideology, individual differences

## Warning Bell: Liberals Implicitly Respond to Group Morality Before Rejecting it Explicitly

Since the widespread reporting of a national “moral values” divide after the 2004 U. S. Presidential election, morality has become a focal point of interest for understanding political ideologies and values. However, the reason liberals and conservatives come to such drastically different moral conclusions on diverse issues such as homosexuality, immigration, and welfare – and why prolonged debate so rarely leads to any kind of synthesis or consensus – is not clear.

One hypothesis is suggested by a participant in an earlier study of the moral intuitions of liberals and conservatives. After listening to a story about two men who engage in consensual anal sex, the participant said:

“I’d have to say it would make me feel weird. That’s my first instinct. Yeah, I mean there’s a weird kind of thing going on in my head there. I feel like there’s a way I’m going to act and a way to act based on this preconceived ideology I have about sex being okay between people, no matter what sex they are. You know, instantly a warning bell comes up, and that bugs me. I quickly discard it; I mean I recognize it and then discard it because it makes me feel silly” (Haidt & Hersh, 2001).

This participant feels an intuitive “warning bell” of discomfort at the thought of gay anal sex, but because such a discomfort conflicts with his explicitly endorsed liberal values, he quickly discards it. One source of partisan differences, then, may be the degree to which liberals and conservatives listen to their gut feelings as a guide to what’s right and what’s wrong.

Building off of explicit moral differences found between liberals and conservatives (Graham, Haidt, & Nosek, 2009), the present studies test a hypothesis suggested by the participant quoted above – namely, that liberals actually share many of the moral intuitions of conservatives, but do not consciously endorse them. We test this idea by focusing on *implicit* aspects of moral judgment – aspects that are relatively automatic, associative, unavailable to introspection, unreportable, and outside of conscious awareness (see Gawronski & Payne, 2010;

Greenwald & Banaji, 1995; Nosek, Hawkins, & Frazier, 2011; and Wilson, 2002 for reviews of the implicit/explicit distinction in attitudes). The studies focus on implicit measurement of moral concerns that are on average consciously endorsed by conservatives, but not by liberals – specifically, those related to loyalty, authority, and sanctity.

### *Moral Foundations Theory*

Employing Shweder, Much, Mahapatra, and Park's (1997) "big three" ethics of autonomy, community, and divinity, Haidt and Hersh (2001) found that liberals primarily endorsed the ethics of autonomy, whereas conservative morality was more likely to include the ethics of community and divinity as well as autonomy. Reviewing anthropological and evolutionary literatures, Haidt and Joseph (2004) expanded Shweder's three ethics, incorporating Fiske's (1992) four relational models into a theory of foundational moral intuitions that different cultures build on to varying degrees. As the theory has developed (Graham & Haidt, 2010, in press; Graham, Haidt, & Rimm-Kaufman, 2008; Haidt & Graham, 2007, 2009; Haidt, Graham, & Joseph, 2009; Haidt, 2012) it describes these five moral foundations: Care/harm, Fairness/cheating, Loyalty/betrayal, Authority/subversion, and Sanctity/degradation. The Care/harm foundation covers basic concerns about the suffering of others, including the compassion and care emphasized by Gilligan (1982). Fairness/cheating covers norms of reciprocal relations, equality, rights, and justice, encompassing the majority of psychological treatments in the Kohlbergian tradition. Loyalty/betrayal covers moral obligations and considerations that come along with group membership, such as expectations of preferential treatment for ingroup vs. outgroup members. Authority/subversion covers the moral obligations of hierarchical relations, such as obedience, respect for superiors, and protection of subordinates. Sanctity/degradation is based on widespread intuitions about divinity based in part on feelings of

moral disgust, but also on transcendent notions of sacredness, elevation, and spiritual purity (see Kass, 1997).

A major goal of Moral Foundations Theory is to describe the ways moral concerns vary across individuals and groups, in order to provide an explanatory account of moral conflicts such as those between political partisans. Haidt and Graham (2007) articulated the hypothesis that the intractability of such moral disputes could be explained by the fact that liberal morality is built primarily upon the Care and Fairness foundations, while conservative morality is built more evenly upon all five foundations.

Graham et al. (2009) found evidence for this prediction using a variety of methods. Foundational moral concerns were measured via self-reported ratings of what is relevant to one's moral judgments (i.e., "Whether or not someone did something to betray his or her group," for Loyalty), agreement/disagreement with more concrete statements (i.e., "People should not do things that are disgusting, even if no one is harmed," for Sanctity), the amount of money one would require (hypothetically) to be willing to do things that violate the moral foundations (e.g., "Kick a dog in the head, hard," for Care), and the frequency of foundation-related word use in sermons of liberal and conservative churches. All of these measures showed the same basic pattern: Liberals were more concerned than conservatives about Care and Fairness, and conservatives were more concerned than liberals about Loyalty, Authority, and Sanctity. Moreover, while conservatives generally endorse all five foundations, liberals strongly endorse Care and Fairness but on average do not endorse the group-focused foundations (for instance, rating them as not very relevant to morality, or disagreeing with statements expressing group-focused moral concerns). This finding has been replicated in nationally-representative U. S. samples (Smith & Vaisey, 2010), several other nations and world areas (Graham, Nosek, Haidt,

Iyer, Koleva, & Ditto, 2011; Van Leeuwen & Park, 2009), and the personal narratives of liberal and conservative Christians (McAdams, Albaugh, Farber, Daniels, Logan, & Olson, 2008).

### *Implicit Morality*

Although ideological differences in the moral foundations have proven robust across multiple samples and methods, all of these methods relied on either self-reported (and consciously endorsed) values, or actions that are the consequence of deliberation and intention such as sermon composition. In the past decade, moral psychology has begun to investigate the automatic, intuitive, and emotional bases of moral judgment (Haidt, 2001), a turn away from the Kohlbergian tradition that tended to treat morality as conscious, deliberative reasoning about dilemmas. Moral judgments have been shown to be affected by manipulations outside the participants' awareness (or at least which the participant would not endorse as justifiable causes of their judgments), such as incidental disgust from smells or messy environments (Schnall, Haidt, Clore, & Jordan, 2008), actions such as cleaning one's hands (Schnall, Benton, & Harvey, 2008; Zhong, Strejcek, & Sivanathan, 2010), and even hypnotic suggestions designed to induce feelings of disgust (Wheatley & Haidt, 2005). Similarly, individual differences in disgust sensitivity have been shown to predict intuitive negativity toward gays, measured both by an Implicit Association Test and by intentionality assessments (Inbar, Pizarro, Knobe, & Bloom, 2009). Neuroimaging studies have also begun to investigate the processes involved in making moral decisions, many of which are not available to conscious introspection (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Luo, Nakic, Wheatley, Richell, Martin, & Blair, 2006; Young, Camprodon, Hauser, Pascual-Leone, & Saxe, 2010).

Recent theoretical and empirical developments of the "new unconscious" movement (Hassin, Uleman, & Bargh, 2005) have transformed not only the study of morality, but of

ideology as well (see Nosek, Graham, & Hawkins, 2010, for a review). For instance, conservatives show greater implicit preferences for high-status groups (straight people, white people, light skin, non-Arabs) over low-status groups (gay people, black people, dark skin, Arabs) than do liberals (Jost et al., 2008). Further, the notion that ideological commitments may be inarticulable or even unavailable to conscious awareness has been a major factor in the revival of ideology as a topic of scientific study by political psychologists (Jost, 2006).

### *The Present Research*

Because explicit political differences in moral foundation endorsement have been robustly demonstrated in several different samples and by several different research teams, they are a promising avenue for demonstrating the phenomenon suggested by the “warning bell” passage above. Given that liberals on average do not endorse the Loyalty, Authority, or Sanctity foundations, the present studies test the hypothesis that liberals will nevertheless implicitly react to stimuli related to these foundations. In short, we predict that liberals will have “warning bells” related to group loyalty, respect for traditions and authorities, and physical/spiritual sanctity that—according to their explicit judgments—they do not heed.

There is some initial evidence for this hypothesis in work on political attitudes and stereotyping. Skitka and Tetlock (1993) showed that liberals feel morally wrong using personal responsibility as a criterion for withholding aid while conservatives do not, and that conservatives are more likely to vindictively punish individuals who deviate from the social order. Skitka, Mullen, Griffin, Hutchinson, and Chamberlin (2002) specified the mechanism of such differences as motivated correction, by which liberals “think twice” and correct their initial default responses of stereotyping, ingroup bias, and personal attributions like victim-blaming. When liberals are tired, distracted or under cognitive load, the authors found, they tend to “act

like” conservatives, who lack the egalitarian motivations to correct their biased intuitive responses to low-status or outgroup members (see also Jost, Kay, & Thorisdottir, 2009, and Eidelman, Crandall, Goodman, & Blanchard, in press, on the automatic nature of aspects of conservative ideology such as system justification and status quo bias). In this paper we apply Moral Foundations Theory as an organizing framework to explain the moral nature of this asymmetry, testing the possibility that liberals may share some of conservatives’ intuitive reactions to Loyalty, Authority, or Sanctity violations, but reject those intuitions when making explicit endorsements.

### Overview of Studies

The present studies test the hypothesis that liberals will show a greater disparity between their implicit and explicit judgments about Loyalty, Authority, and Sanctity than will conservatives. Specifically, liberal implicit reactions to violations of these foundations will “look” more conservative than what is suggested by their explicit judgments and beliefs. We do not predict any such discrepancy for conservatives.

We note, however, that when an intuition is suppressed regularly enough to become a habit, then this suppression can itself become automatized (see e.g. Moskowitz, Gollwitzer, Wasel, & Schaal, 1999), so it could be hard to determine whether ideological suppression (or “political correction”) is occurring. For this reason, the studies take a multimethod approach in order to measure moral reactions in several different ways. From self-reported gut reactions (Study 1) to neural responses (Study 4), the studies are intended to cover a wide swath of the methods available in implicit social cognition and social neuroscience. All studies rely on within-subject or between-subjects experimental effects to demonstrate differences in reactions to

specific foundation-related stimuli. The measures vary in richness of moral content provided, from simple words (Study 2) to pictures (Study 3) to sentences (Studies 1 and 4).

### Study 1

In Study 1 participants were randomly assigned either to complete the Moral Foundations Questionnaire normally (reporting their consciously-endorsed values and judgments), or to give their first immediate “gut” reactions. Although self-reported gut reactions are a direct measure of attitudes, they show more convergence with implicit measures than do self-reports of “actual” feelings (Ranganath, Smith, & Nosek, 2008). In the gut reaction condition, participants were also given the opportunity to indicate their “real” endorsed answers alongside their self-reported gut reactions. Following Ranganath et al. (2008), this second rating was included to help clarify for participants that their initial responses could be different from their more considered judgments.

### Method

#### *Participants*

Participants were 842 adult visitors (337 female, median age 31) to YourMorals.org. All participants had previously provided demographic information at registration, including a 10-option political ideology self-report item (strongly liberal, liberal, slightly liberal, moderate, slightly conservative, conservative, strongly conservative, libertarian, don't know/not political, other). There were 512 liberals, 72 moderates, 93 conservatives, and 68 libertarians.

#### *Procedure*

All participants self-selected to take the Moral Foundations Questionnaire (MFQ; Graham et al., 2011). After indicating their consent, participants were randomly assigned to one of two conditions. Because the conditions rely on reading the instructions, an Instructional Manipulation Check was used to rehabilitate those who didn't read the instructions the first time.

Participants read a block of text that explained that although there was a “continue” button at the bottom of the page, they should instead click on the “Instructions” heading at the top of the screen. Participants were told that this action would demonstrate to the researchers that study participants were reading the instructions properly. Participants were asked to read the instructions again if they clicked on the “continue” button and repeated this procedure until they completed the assignment correctly by clicking on the “Instructions” button (see Oppenheimer, Meyvis, & Davidenko, 2009). Five hundred one participants, as instructed, never clicked the “continue” button, 292 clicked it once, 39 clicked it twice, 9 clicked it three times, and one clicked it five times.<sup>1</sup>

Instructions for the two parts of the MFQ were as follows [additional instructions for gut condition in brackets]:

Relevance section: When you decide whether something is morally right or wrong, to what extent are the following considerations relevant to your thinking? [**Please indicate your first intuitive “gut” reaction, then indicate your final reasoned-out answer.**]

Judgments section: Please indicate your agreement or disagreement with the following statements. [**Please indicate your first intuitive “gut” reaction, then indicate your final reasoned-out answer.**]

In addition, next to the six response options for each question (“not at all relevant” to “extremely relevant” and “strongly disagree” to “strongly agree”), the phrases “**First gut reaction:**” and “**Final reasoned-out answer:**” were added in the gut condition. After filling out both parts of the MFQ, participants in both conditions were taken to a debriefing screen that gave participants feedback about their moral foundations scores and explained the purpose of the scale and manipulation.

## Results

---

<sup>1</sup> Results are reported for all participants; the same pattern of results was found when we included only those who passed this manipulation check the first time.

*Ideological differences remain in self-reported gut reactions.* Figure 1 shows the relations between ideology and each foundation average. As expected, the control condition replicated the ideological differences found previously (Graham et al., 2009): conservatives endorsed Care and Fairness concerns less, and group-focused concerns more, than liberals did, average  $|r| = .45$ ,  $.31 \leq |r|s \leq .56$ ,  $ps < .001$ . This pattern was also found in the experimental condition, for both self-reported gut reactions (shown in Figure 1), average  $|r| = .41$ ,  $.22 \leq |r|s \leq .54$ ,  $ps < .01$ , and final reasoned-out answers, average  $|r| = .40$ ,  $.21 \leq |r|s \leq .56$ ,  $ps < .01$ .

*Across the ideological spectrum, reports of first gut reactions were no different than reports of endorsed moral opinion.* Overall effects of condition (gut vs. control) were found for Loyalty ( $t = -2.50$ ,  $p = .01$ ,  $d = .17$ ), and marginally for group-focused concerns generally ( $t = -1.70$ ,  $p = .09$ ,  $d = .12$ ), such that these concerns were endorsed more in the gut condition than in control. However, multiple regressions predicting foundation averages from condition, ideology, and their interaction revealed no effects of condition, and no condition-ideology interaction, for any of the foundations. These regressions all showed significant effects of ideology, average  $|\beta| = .43$ ,  $.27 \leq |\beta|s \leq .55$ ,  $ps < .0001$ , in the directions shown in Figure 1.

*Within the experimental condition, differences between first-gut-reaction and final-answer ratings were consistently small regardless of ideology.* Paired-samples t-tests revealed significant differences between self-reported gut and final answers among those in the experimental condition for all five foundation averages, average  $|t| = 3.89$ ,  $2.47 \leq |t|s \leq 8.21$ ,  $ps \leq .014$ . For all foundation averages the gut ratings were higher than the final answer ratings, with the exception of Authority, for which the difference was in the opposite direction. The gut-final differences, however, were very small, ranging from  $-.07$  to  $.22$ , with an average of just  $.11$  on the 6-point scale shown in Figure 1. The gut-final difference varied with politics for Fairness,  $r =$

-11,  $p = .049$ , but did not vary by politics for the other four foundations. In general, there were only very small distinctions made between first gut reactions and final reasoned-out answers on the MFQ, even when these were answered separately side by side.

*Liberals reported that their first gut reactions were contrary to group-focused moral concerns.* As Figure 1 illustrates, liberals on average used the bottom half of the scale responses (indicating little moral relevance and disagreeing with statements expressing foundation-related concerns) when completing the MFQ normally, *and* when indicating their first gut intuitions. Follow-up analyses confirmed that liberals in the gut condition were significantly below the scale midpoint of 2.5 for their reports of group-focused gut intuitions,  $M=1.86$ ,  $t(234) = -12.40$ ,  $p < .001$ ,  $d = .81$ .<sup>2</sup> In contrast, moderates' group-focused gut answers were not different from the midpoint,  $M=2.67$ ,  $t(28) = 1.39$ , *n.s.*, and conservative gut answers were above the midpoint,  $M=3.13$ ,  $t(44) = 5.11$ ,  $p < .001$ ,  $d = .76$ , indicating that their self-reported first gut reactions were in support of the group-focused foundations, not opposed to them.

### Discussion

When participants were asked to self-report their first intuitive “gut” reactions, they gave the same patterns of foundation endorsements as when simply indicating their endorsed moral opinions – with the same strong effects of ideology on foundation endorsement found in previous studies (Graham et al., 2009; 2011). Even when asked to give side-by-side separate answers for their first gut reaction and final reasoned-out answer, participants indicated that these were largely the same, deviating by only a tenth of a scale point on average. Thus, according to participants' self-reports, there are *no* implicit-explicit differences in foundation-related moral judgments and concerns. Specifically, liberals reported that their first gut reactions are in line

---

<sup>2</sup> Differences were statistically significant ( $ps < .001$ ) for Loyalty, Authority, and Sanctity separately as well. Care-Fairness gut reports were above the scale midpoint for all political groups,  $ps < .001$ ,  $ds > 1.00$ .

with Care and Fairness concerns, but are opposed to Loyalty, Authority, and Sanctity concerns. It is important to note, however, that these are *self-reports* of gut reactions, which may be quite different from actual implicit reactions. Study 1 suggests that across the political spectrum, people think that their intuitive gut reactions are the same as their consciously endorsed moral opinions. The purpose of Studies 2 through 4 is to test whether – and for whom – this is correct.

### Study 2

Unlike Study 1, Study 2 bypassed self-report altogether in its search for implicit reactions to moral foundation-related stimuli. Adapting an evaluative priming procedure used by Ferguson (2007), this study investigated reaction times to assess automatic negativity associated with foundation-specific vice words. Ferguson (2007) found that automatic associations of positivity with the word “equal” predicted scores on a measure of subtle prejudice, and suggested that such associations reflected stable (and morally-relevant) automatic goals to be egalitarian. We modified this procedure to include words related to all five foundations, and used vice instead of virtue words, with the expectation that these would yield more reliable priming effects since bad is stronger than good (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Rozin & Royzman, 2001). Participants explicitly rated the vice words as well.

### Method

#### *Participants*

Participants were 279 students at the University of Virginia (181 female) who participated in the lab study in exchange for partial course credit or payment of \$8. In pretesting all participants had previously recorded their demographic information, including a 7-point politics item (strongly liberal to strongly conservative). There were 138 liberals (very liberal to

slightly liberal), 55 moderates, 79 conservatives (slightly conservative to very conservative), and 7 did not answer.

### *Procedure*

Participants first explicitly rated 40 foundation-related words (shown in Supplemental Appendix A) on a scale from 1 (very negative) to 7 (very positive). Order of words was randomized for each participant. Then, pairs of words were presented sequentially; participants were told to ignore the first word (prime) and evaluate the second word (target) as positive or negative as quickly as possible. For simplicity, the study design contrasted two groups of vice primes, one set for Care-Fairness (hurt, cruel, cheat) and one set for Loyalty-Authority-Sanctity (traitor, revolt, sin); there were also 13 control primes (neutral words, e.g., chair, shelf). Foundation primes were selected for face validity and simplicity from a list of foundation words used in a previous text analysis study (Graham et al., 2009, Study 4). All primes and target adjectives are shown in Supplemental Appendix A. There were four practice trials and 48 test trials, in which every foundation prime word preceded two negative and two positive words. The prime appeared for 150ms, followed by 100ms of blank screen for a total 250ms stimulus onset asynchrony (SOA). SOA and other specifics of the evaluative priming procedure were adapted from Ferguson (2007). The target adjective then appeared until the participant indicated positive or negative by pressing specially-marked keys on the keyboard. The intertrial interval was 2000ms. Order of primes and targets was randomized for each participant. Participants were then given a Boggle distractor task, several questionnaires and an evaluative conditioning task that are not relevant for the current report and are not discussed further.

### Results

*Explicit word ratings mirrored previous political differences in moral foundation endorsement.* As shown in Figure 2, explicit word ratings showed the same basic political effects found in Study 1: liberals showed a greater explicit endorsement (operationalized as higher ratings of virtue than vice words) for Care and Fairness than conservatives, who showed greater endorsement of the three group-focused foundations than liberals did. Political conservatism predicted higher ratings of Loyalty-Authority-Sanctity virtue words ( $\beta = .25, p < .001$ ) and Care-Fairness vice words ( $\beta = .16, p < .01$ ), and lower ratings of Loyalty-Authority-Sanctity vice words ( $\beta = -.22, p < .001$ ); it did not predict ratings of Care-Fairness virtue words ( $\beta = .03, p = .65$ ). Examining only the six vice words used in the priming procedure, political conservatism predicted lower self-reported ratings of the Loyalty-Authority-Sanctity vice primes (traitor, revolt, sin;  $\beta = -.29, p < .001$ ) but did not predict ratings of the Care-Fairness vice primes (hurt, cruel, cheat;  $\beta = .01, p = .83$ ).

*Liberals show as much automatic negativity to violations of Loyalty-Authority-Sanctity as to violations of Care-Fairness.* Automatic positivity/negativity toward the foundations was calculated as follows, based on Ferguson (2007). Error trials, very slow trials ( $>2000\text{ms}$ ) and very fast trials ( $<250\text{ms}$ ) were excluded from analyses. Positive facilitation scores were created for each foundation by subtracting the mean response time for positive targets following control primes from response times for positive targets following foundation primes. Negative facilitation scores were calculated the same way, comparing response times for negative targets following control primes to response times for negative targets following foundation primes. A main index of positivity toward each foundation was calculated by subtracting the negative facilitation score for that foundation from the positive facilitation score for that foundation. Thus each participant produced two main scores, one for Care-Fairness (CF) and one for Loyalty-

Authority-Sanctity (LAS). (Automatic positivity D-scores were also calculated separately for CF primes, LAS primes, and control primes, by subtracting response times for positive targets from response times for negative targets and dividing by the standard deviations; results showed the same patterns.)

Significant negative priming effects were found for both the CF vice primes ( $t(272) = -2.89, p < .01, d = 0.17$ ) and LAS vice primes ( $t(271) = -3.21, p < .001, d = 0.19$ ), meaning that participants overall showed automatic negativity to vice words of both foundation groups. Political ideology did not predict differences in automatic negativity toward Care-Fairness vices ( $\beta = .03, p = .61$ ), and it did not predict differences in automatic negativity toward Loyalty-Authority-Sanctity vices ( $\beta = .00, p = .94$ ). Examined in isolation, liberal participants showed significant negative priming effects for both CF vice primes ( $t(136) = -2.16, p = .032, d = 0.19$ ) and LAS vice primes ( $t(136) = -2.43, p = .017, d = 0.21$ ), with no difference in magnitude between the two ( $t(136) = -0.16, ns$ ).

In aggregate analyses, explicit and implicit reactions to the same six vice words were standardized and treated as repeated measures, with political ideology as the predictor. These analyses showed no interaction of politics with implicit/explicit for CF, but a significant interaction of politics with implicit/explicit for LAS ( $F(256) = 3.13, p = .006, d = 0.11$ ), indicating significantly different slopes across ideology for the implicit and explicit measures of group-focused foundations.<sup>3</sup>

## Discussion

Study 2 provides the first evidence of automatic negativity to foundation-related stimuli not captured by self-report: Liberals and conservatives report different opinions about these

---

<sup>3</sup> The same pattern of results was found when using the full explicit indices (ratings of all virtue words minus ratings of all vice words) in these analyses.

foundation-related vice words, but show the same automatic reactions to them. In the case of the Care-Fairness primes (hurt, cruel, cheat), it's possible that no differences were found in automatic negativity because there were no political differences of opinion about these words even explicitly. However, this explanation does not hold for the Loyalty-Authority-Sanctity primes: although liberals explicitly evaluated the words traitor, revolt, and sin less unfavorably than did conservatives, they showed significant automatic negativity to these words that was just as strong as that of conservatives – moreover, their automatic negativity toward these words was just as strong as their automatic negativity to the Care-Fairness vice words. This study provides the first evidence of an asymmetry in implicit-explicit discrepancies across the political spectrum: Conservative priming effects were in line with their self-report ratings, but liberal priming effects suggested intuitive reactions to these words that are at variance with their self-reports.

One drawback of Study 2 is that it relied on single words, without any context. It is possible that the automatic negativity found toward the vice primes simply reflects the valence of the words in English, without any moral intuitions involved – that is, perhaps liberals automatically categorize “sin” as a negatively-valenced word, and perhaps they have a slight flash of negative affect to it, in the way that Zajonc (1980) described. Yet they still might lack conservatives' intuitive responses to signs of sin in the world. Further, even though significant priming effects were found, the relative unreliability of priming effects could have limited our ability to find relationships between it and individual-difference variables such as political ideology. Study 3 uses a different paradigm and different kinds of stimuli in order to address these issues.

### Study 3

Study 3 attempted to replicate the findings of Study 2 using a different method. To do this we turned to the Affect Misattribution Procedure, an implicit measure shown to have relatively strong and reliable effects (AMP; Payne, Cheng, Govorun, & Stewart, 2005). The AMP instructs users to ignore a prime photo and rate a Chinese character as either less or more pleasant than average according to their gut feeling about it; the idea is that even when instructed to ignore it, participants will misattribute their affective reactions produced by the prime onto the character. This method also allowed us to use foundation-related pictures instead of just single words, allowing for more moral richness in the stimuli used.

### Method

#### *Participants*

Participants were 106 students at the University of Virginia (58 female) who participated in the lab study in exchange for partial course credit. As in Study 2, in pretesting all participants had previously recorded their demographic information, including a 7-point politics item (strongly liberal to strongly conservative). There were 48 liberals (very liberal to slightly liberal), 24 moderates, 33 conservatives (slightly conservative to very conservative), and one who did not answer.

#### *Procedure*

A screen informed participants that they would see two pictures, a photo followed by a Chinese character, and that they would be asked to rate whether the Chinese character was more or less pleasant than the average Chinese character. They were also given the following instructions: “Please note that sometimes the photos flashed prior to the characters can influence people’s ratings of the Chinese characters. Please try your best not to be influenced by the photographs. Instead, please give us an honest judgment of how pleasant or unpleasant is your

reaction to each Chinese character. Of course, there are no right or wrong answers. Just report your ‘gut reaction.’” The study design contrasted photos representing virtues and vices of Care, Fairness, Loyalty, Authority, and Sanctity; all 40 foundation-related pictures are shown in Supplemental Appendix B. Pictures were selected from a set of over 200 photos (compiled by the author and five research assistants) based on three criteria: 1. face-valid relations to one foundation, 2. relative lack of relation to other foundations, and 3. relative simplicity. The prime pictures appeared for 150ms, followed by 100ms of blank screen and then 100ms of the Chinese character. The participants then had to press a specially-marked key indicating either “less pleasant” (than the average Chinese character) or “more pleasant.” Each participant saw each foundation-related picture twice, for a total of 80 trials. Order of primes and Chinese characters was randomized for each participant. Participants were also given a Boggle distractor task, questionnaires and a priming task<sup>4</sup> that are not relevant for the current report and not discussed further. Finally, they were debriefed on the design and aims of the study.

## Results

*Liberals gave weaker explicit ratings of Loyalty-Authority-Sanctity pictures than did conservatives.* In a separate data collection, 45 students at the same university explicitly rated the moral content of 38 of the 40 pictures used in the study, on a scale ranging from 1=strongly morally bad to 7=strongly morally good. Political ideology had been reported previously using the same 7-point scale described above. Pictures were presented in randomized order, and averages for each foundation were created by subtracting the ratings of the vice pictures from the ratings of the virtue pictures. There were no political differences in ratings of Care ( $r = -.01$ ,  $p = .94$ ) or Fairness pictures ( $r = -.08$ ,  $p = .63$ ). However, liberals gave weaker explicit ratings than

---

<sup>4</sup> Order of AMP and priming task was counterbalanced, but task order had no effects on any of the variables of interest.

conservatives for pictures representing Loyalty ( $r = .38, p = .01$ ), Authority ( $r = .23, p = .13$ ), and Sanctity ( $r = .55, p < .001$ ). Liberals explicitly reported that the Loyalty-Authority-Sanctity virtue pictures were less morally good ( $.13 \leq rs \leq .49$ ) – and vice pictures less morally bad ( $-.50 \leq rs \leq -.18$ ) – than did conservatives. AMP analyses thus concentrate on whether liberals were also less implicitly reactive to these pictures than conservatives.

*Across foundations, positive affect was misattributed from virtue pictures to targets, and negative affect from vice pictures to targets.* Following Payne et al. (2005), automatic reactions to foundation-related prime pictures as measured by the AMP were based on how many times the Chinese character was rated as more pleasant (vs. less pleasant) than average when it followed the picture. Responses were either “less pleasant”=0 or “more pleasant”=1. For each person, each picture had a rating, either 0, 0.5 or 1, based on averaging their two responses when the picture was the prime. Averaged across people, picture scores significantly less than .5 indicate a negative affective response to the prime, greater than .5 indicate positive affective response, and .5 is chance (no effect), assuming no mean tendency to give either response more often than the other regardless of the stimuli. For 31 of the 40 pictures, mean scores were significantly different from .5 ( $ps < .01$ ), always in the expected direction (“more pleasant” selected more often than chance following virtue pictures, “less pleasant” selected more often than chance following vice pictures).<sup>5</sup> Every foundation’s set of virtue pictures led to responses of “more pleasant” greater than chance ( $ts(103) > 5.00, ps < .001, ds > .49$ ), and every foundation’s set of vice pictures led to responses of “less pleasant” greater than chance ( $ts(103) > 3.00, ps < .01, ds > .30$ ), with the exception of Loyalty vice pictures, which trended in the

---

<sup>5</sup> In a heartening side finding with one of the nine exceptions, the Virginia Tech logo (used as one of the Loyalty vice images, as it is UVA’s rival) was a significant negative prime for the first part of the study, but during the study there was a tragic shooting at Virginia Tech, and there was an outpouring of support for VT on the UVA campus. After the shooting, the VT logo was no longer a significant negative prime for UVA students.

expected direction but not significantly ( $t(103) = 1.30, p = .19$ ). A general AMP effect index was created for each foundation by subtracting their vice average from their virtue average; these indices were significantly greater than 0 for all five foundations ( $ts(103) > 5.00, ps < .001, ds > .49$ ). In general, the foundation-related pictures produced strong and reliable AMP effects.

*Liberals showed implicit reactions to Loyalty-Authority-Sanctity pictures as strongly as did conservatives.* Regression analyses predicting each foundation's AMP effect index from political ideology showed that liberals were more reactive to Fairness pictures than were conservatives ( $\beta = -.21, p = .03$ ). In other words, Fairness photographs had a bigger effect on liberal ratings of Chinese characters than on conservative ratings of those characters. There were no effects of politics on the other four foundation AMP effect indices, or on an averaged Loyalty-Authority-Sanctity AMP effect index ( $|\beta|s \leq .11, ps > .28$ ). Further analysis revealed politics effects on Fairness virtue pictures ( $\beta = -.31, p < .01$ ) but not on Fairness vice pictures ( $\beta = .04, p = .66$ ). Within the four Fairness virtue pictures (see Supplemental Appendix B), politics effects were in the same direction for all four pictures, but were significant only for the picture of Black women marching for equal pay and housing rights ( $\beta = -.25, p = .01$ ).

Examined in isolation, liberals showed a marginal AMP effect for Care ( $t(46) = 1.72, p = .09, d = .25$ ) and a significant AMP effect for Fairness ( $t(46) = 3.99, p < .0001, d = .58$ ). Most notably, liberals also showed significant AMP effects for Loyalty ( $t(46) = 5.46, p < .0001, d = .80$ ), Authority ( $t(46) = 7.01, p < .0001, d = 1.02$ ), and Sanctity ( $t(46) = 2.64, p = .011, d = .39$ ).

### Discussion

For automatic reactions to the group-focused foundations, the AMP findings replicated those of Study 2: liberals again showed significant automatic affective reactions to Loyalty, Authority, and Sanctity, just as strongly as conservatives did. These liberal reactions were also

on average *stronger* than liberal reactions to Care and Fairness; the strongest AMP effects for liberals were with the Authority and Loyalty pictures. This is particularly surprising in light of the fact that liberals' explicit ratings of the LAS pictures were weaker (for both virtue and vice pictures) than conservatives' ratings. The results for Fairness provided evidence of differences in affect misattribution, in that liberals had more implicit reactivity than conservatives did. This political difference was specific to Fairness virtue pictures: participants across the political spectrum showed automatic positivity to these pictures (misattributing positive affect to the Chinese characters that followed them), but liberal participants showed this effect more than conservatives. The fact that automatic political differences were found with Fairness but not the other foundations is particularly surprising in light of the fact that self-reported political differences tend to be highest for Authority and Sanctity, not Fairness (see Graham et al., 2011, for a replication of this finding across several nations and world areas).

Although the pictures used in Study 3 provide richer moral content and context than the single words used in Study 2, they could be *too* rich for implicit methods; there are a variety of cognitions that could be primed by the pictures, and ways in which they could be interpreted. In Study 4 we addressed this by using entire sentences giving moral opinions, increasing both moral richness and clarity of the moral content of the stimuli.

#### Study 4

Study 4 continued the systematic increasing of richness of moral content, using entire sentences expressing moral opinions related to foundational concerns. Like Studies 2 and 3, Study 4 sought to assess moral reactions outside of self-report, this time using neural activity as measured by electroencephalographic (EEG) recordings of event-related potentials (ERPs). Recent work in psycholinguistics has shown that meaning extraction and sense-making processes

occur rapidly and continually during reading, and that neural reactions to specific words can be recorded even before the entire statement is finished (Van Berkum, 2008). This study was inspired by a recent application of linguistic ERP methodologies to morally offensive statements (Van Berkum, Holleman, Nieuwland, Otten, & Murre, 2009), in which two ERP components were used to gauge automatic detection of violations with participants' moral values. The first component, the N400, has been shown to track semantic incongruities and linguistic violations of both surface expectations ("He took his coffee with cream and *dog*"; Kutas & Hillyard, 1980, 1984) and deeply-held moral values ("I think euthanasia is an *acceptable* course of action"; Van Berkum et al., 2009). N400 potentials have also been used to study immediate reactions to strongly-valenced political stimuli (Morris, Squires, Taber, & Lodge, 2003). The second component, the late positive potential (LPP), is associated with intensive affective processing, particularly for negatively-valenced stimuli (Cacioppo, Crites, Gardner, & Berntson, 1994; Ito, Larsen, Smith, & Cacioppo, 1998). Following Van Berkum et al. (2009), we recruited partisans on either end of the political spectrum and constructed sentences on which their explicit answers were likely to differ. ERP components allow us to evaluate with great precision the time-course of immediate reactions to critical words in the morality sentences, and test – using neurophysiological as well as behavioral data – whether partisans' implicit reactions differ from their conscious endorsements of the statements. The N400, as an indicator of conflict detection, and the LPP, as an indicator of intensive affective processing, are particularly well-suited to testing our hypothesis that liberals have immediate affective reactions to group morality concepts that may oppose their ideologies and produce corrective efforts to bring them in line.

## Method

### *Participants*

Participants were 65 students (42 female) at the University of Virginia. Students self-identifying as very liberal, liberal, conservative, or very conservative during pretesting were invited to participate in the study in exchange for partial course credit. Students self-identifying as slightly liberal, moderate, or slightly conservative were not eligible to sign up for the study. There were 39 liberals and 26 conservatives.

### *Materials and Procedure*

Electroencephalographic recordings were made using the Active Two system (BioSemi, Amsterdam, The Netherlands). A stretch lycra cap was placed on participants' heads and 32 Ag/AgCl-tipped electrodes arranged according to the 10/20 international labeling system were attached to the cap. A small amount of electrolyte gel was applied to each electrode to allow for low electrode offsets. Additional electrodes were placed at the infra-orbital sites of both eyes to monitor for vertical and horizontal eye movements, and at the left and right mastoids behind the ears to provide reference points where there should be no electrical signals. Data were sampled at 500 Hz, with a feedback loop between a Common Mode Sense active electrode and a Driven Right Leg passive electrode forming the ground (so that all electrical signals collected are relative to these two electrodes).

Once participants were fitted with the EEG cap and electrodes, they were seated in front of a computer, shown the camera above the monitor, and told to limit head movements and only use the arrow keys on the keyboard. They were then instructed on the screen that they would see a series of sentences one word at a time, and after each sentence they were to use the arrow keys to select from four options: Strongly disagree, Slightly disagree, Slightly agree, and Strongly agree. Each word was displayed for 750ms, with a 1000ms interval between participant answer and the start of the next sentence.

Each sentence had a critical word indicating the direction of moral opinion about the foundation-related moral issue. For instance, in the sentence “Total equality in the workplace is necessary,” the word “necessary” is the critical word during and after which ERPs were examined. Sixty-six different sentences were written supporting either Care-Fairness or Loyalty-Authority-Sanctity; for each sentence there were four possible variations, of which each participant saw two. The critical word either supported (“necessary”) or rejected (“unrealistic”) the foundational concerns, allowing us to compare ERPs for agreement or disagreement. As an extra control, the critical word appeared either at the end of the sentence (as in the example above) or at the beginning (“It is unrealistic to have total equality in the workplace”), allowing us to compare reactions to the same words when their semantic moral content was or was not established. All sentences are shown in Supplemental Appendix C. Sentences were generated by the first author and four research assistants, then narrowed down and modified by the first author. Within each block of four sentences, each participant either saw the first sentence and later in the study saw the third sentence, or saw the second sentence and later saw the fourth sentence (this was to avoid participants recognizing sentences in reversed order or switched critical words). Each participant was randomly assigned to one of two conditions containing a specific set of sentences in a scrambled order (with the above order parameters met). Each participant saw a total of 132 sentences, with a break after the first 66. After completing the EEG recording portion of the study, participants removed their EEG caps, filled out the Moral Foundations Questionnaire on paper, and went through debriefing (and some hair-washing).

## Results

*Liberals and conservatives showed opposite patterns of agreement for sentences supporting and rejecting foundation-related values.* Because the sentences were intended to elicit

different patterns of agreement and disagreement for liberals and conservatives, we expected that the sentences supporting Care-Fairness and rejecting Loyalty-Authority-Sanctity would be met with agreement by liberals and with disagreement by conservatives. This pattern was not expected to differ based on whether the critical words appeared at the beginning or end of the sentences. Figure 3 shows the predicted pattern of results, for sentences where the critical word appeared at the end (left panel) and those where it appeared at the beginning (right panel). In both cases, liberals tended to agree with pro-CF and anti-LAS sentences, and disagree with anti-CF and pro-LAS sentences, and conservatives showed the opposite pattern. Liberals showed significant agreement/disagreement with sentence types in the predicted directions for all four sentence types ( $|t|s(36) > 7.30$ ,  $ps < .0001$ ,  $ds > 1.20$ ), and conservatives showed significant agreement/disagreement in opposite directions ( $|t|s(25) > 2.40$ ,  $ps < .05$ ,  $ds > .47$ ), with the exception of pro-LAS sentences, with which conservatives showed only marginal agreement ( $t(25) = 1.86$ ,  $p = .07$ ,  $d = .36$ ).

*Preprocessing and analysis strategy for EEG data.* EEG procedures and analysis were adapted from Van Berkum et al. (2009) and Kutas and Hillyard (1980; 1984). Although Van Berkum et al. (2009) examined only the sentences to which participants strongly agreed or strongly disagreed, all trials were examined in this study, because of the interest in looking for neural reactions to moral value incongruities even in cases where no strong reaction is explicitly reported. All EEG data were processed using MATLAB and EEGLAB (Delorme & Makeig, 2004), re-referenced to the linked mastoids, and bandpass filtered at .1 to 15 Hz. Every trial for every participant was visually inspected for eye movements, and trials containing blinks or other abnormalities were removed from analyses. EEG data from 17 participants were discarded because more than 30% of their trials were problematic. ERP averages were created around an

epoch of 200ms before critical word appearance to 1000ms after. After manual rejection of eyeblink and other problematic trials, LPP and N400 averages were created for each of 8 triggers, 8 types of critical word events: 1. critical word supporting CF, end of sentence, 2. critical word supporting CF, beginning of sentence, 3. critical word rejecting CF, end of sentence, 4. critical word rejecting CF, beginning of sentence, 5. critical word supporting LAS, end of sentence, 6. critical word supporting LAS, beginning of sentence, 7. critical word rejecting LAS, end of sentence, 8. critical word rejecting LAS, beginning of sentence.

Only one published study (Van Berkum et al., 2009) has used ERPs to gauge moral intuitions, and Study 4 represents the first attempt to study moral foundations using ERPs in response to sentences shown one word at a time. Because of these exploratory facets of the study, it was important to carefully verify the existence of N400 and LPP signatures before moving to foundation or ideological group comparisons. Therefore, we used the most stringent control, comparing reactions to the same critical words at different points in the sentences. The analysis strategy focused on visual inspection of the averaged waveforms for the two groups (liberals and conservatives), comparing reactions to the same critical words at the beginning and at the end of the sentence (i.e., trigger 1 compared to trigger 2) to search for event-related processing (N400 or LPP) specific to the moral semantic content represented by the appearance of the value-specifying word in non-control trials. Both ERPs were examined using seven electrodes on the central-parietal scalp region (centered on Pz, and also including Cp1, Cp2, P3, P4, PO3, and PO4). Based on past literature (Cacioppo et al., 1994; Kutas & Hillyard, 1980, 1984; Van Berkum et al., 2009), peak maximums between 500ms and 800ms were captured to

gauge the LPP,<sup>6</sup> and peak minimums between 300ms and 500ms were captured to gauge the N400, for each trigger type at each electrode for each participant. Peak signals greater than 75  $\mu\text{V}$  or less than  $-75 \mu\text{V}$  (0.4% of all trials) were not included in peak analyses. Once presence of N400 or LPP was established in comparisons to control trials, the next analyses centered on comparisons of foundation-congruent and foundation-incongruent trials (e.g., pro-CF vs. anti-CF). Because they involved different semantic content and different critical words, responses to CF sentences and LAS sentences were examined separately.

*Anti-Care-Fairness endorsements elicited N400 responses in liberals, but not in conservatives.* Figure 4 shows comparisons of the four types of Care-Fairness-related trials, for liberals (left panels) and conservatives (right panels). Each panel compares the trials during critical words at the end of the sentence (bold lines) to trials during the same critical words at the beginning of the sentences (thin lines). The top two panels show responses to Pro-CF sentences (e.g., “People who lobby for animal rights are *admirable*.”/“*Admirable* people lobby for animal rights.”). For both ideological groups, there were no signs of N400 or LPP potentials differentiating trials with semantic content from control trials. For the Anti-CF sentences (“People who lobby for animal rights are *naïve*.”/“*Naïve* people lobby for animal rights.”), there was a strong (10  $\mu\text{V}$ ) N400 signature for liberals centered at 400ms after stimulus onset when the critical words appeared at the end of the sentence, but not when those same critical words appeared at the beginning of the sentence (see bottom left panel). However, analyses of liberals’ peak minimums between 300ms and 500ms revealed no significant difference between critical and control trials,  $t(18) = 1.35, p = .19$ . Comparisons of CF-congruent and CF-incongruent test trials (words at end of sentence only) are shown together in Figure 5 (note that these are the same

---

<sup>6</sup> Although Cacioppo et al. (1994) and Ito et al. (1998) went to 900ms in their LPP windows, in this study the next word appeared 750ms after the critical word appeared, and 900ms ran into early event-related potentials related to seeing that next word.

averaged waveforms shown in Figure 4, now shown together without control trials). While conservative intuitive reactions to Pro-CF and Anti-CF endorsements were very similar, liberals showed more reactivity to the Anti-CF statements than the Pro-CF statements, specifically in the N400 negative potential at 400ms. What looks to be a late positive potential (LPP) around 600-800ms in this comparison, however, is shown to be no different from control trials in Figure 4 (bottom left panel).

*Pro-Loyalty-Authority-Sanctity endorsements elicited LPP responses in liberals, but not in conservatives; Anti-LAS endorsements elicited LPP responses in both conservatives and liberals.* Figure 6 shows comparisons of the four types of Loyalty-Authority-Sanctity-related trials, for liberals (left panels) and conservatives (right panels). Again, each panel compares the trials during critical words at the end of the sentence (bold lines) to trials during the same critical words at the beginning of the sentences (thin lines). The top two panels show responses to Pro-LAS sentences (e.g., “Fulfilling the traditional duties of your gender is *noble*.”/“It is *noble* to fulfill the traditional duties of your gender.”). For conservatives there was no sign of event-related reactions to Pro-LAS endorsements, but for liberals there was a 10- $\mu$ V late positive potential centered around 600ms in response to these Pro-LAS statements. (As noted above, the LPP is associated with intensive affective processing of negatively-valenced stimuli.)

Analyses of liberals’ peak maximums between 500ms and 800ms confirmed that the Pro-LAS end-of-sentence trials had higher peaks than the beginning-of-sentence trials,  $t(20) = 2.06$ ,  $p = .053$ ,  $d = 0.45$ . For the Anti-LAS sentences (“Fulfilling the traditional duties of your gender is *outdated*.”/“It is *outdated* to fulfill the traditional duties of your gender”) the waveforms showed signs of LPP responses for both conservatives *and* liberals. Analyses of conservatives’ peak maximums between 500ms and 800ms confirmed that the Anti-LAS end-of-sentence trials had

higher peaks than the beginning-of-sentence trials,  $t(19) = 2.48$ ,  $p = .023$ ,  $d = 0.55$ . And despite the fact that they explicitly *agreed* with the Anti-LAS statements, liberals' peak maximums also confirmed that the Anti-LAS end-of-sentence trials had higher peaks than the beginning-of-sentence trials,  $t(20) = 2.87$ ,  $p = .009$ ,  $d = 0.63$ . Comparisons of LAS-congruent and LAS-incongruent test trials (all critical words at end of sentence) are shown together in Figure 11 (note again that these are the same averaged waveforms shown in Figure 10, now shown together without control trials). Of special note in this figure is the fact that for liberals, the LPP response to Pro-LAS endorsements is stronger than the response to Anti-LAS endorsements, while for conservatives the pattern is the exact opposite.

### Discussion

Study 4 provides further evidence for our hypothesis – that liberals have reactions to group-focused moral concerns that are not consciously endorsed – using event-related brain potentials. A particular strength of the design lies in the precision of the control trials, presenting the same critical words but presenting them before the moral opinion expressed in them becomes clear. A further strength lies in the use of EEG, a very different method from the affective priming techniques used in Studies 2 and 3.

Different patterns of neural reactions were found for liberals and conservatives, including liberals' apparent N400 response to Anti-CF statements and LPP response to Pro-LAS statements, neither of which conservatives showed. Another difference was in the relative magnitude of the LAS LPP responses, with liberals more reactive to Pro-LAS values and conservatives more reactive to Anti-LAS values. These differences reflect the differences shown in Study 1 and in previous research, and are in line with participants' patterns of explicit agreement and disagreement with the sentences themselves.

However, this study also provided evidence of neural patterns associated with affective reactions to negative stimuli (Ito et al., 1998) in liberals to statements rejecting group-focused moral values. This is particularly striking because in their explicit reactions to these statements, they were more likely to agree with these statements than disagree. This finding builds on the findings of Studies 2 and 3 to suggest that moral intuitions and explicitly-endorsed moral values differ in liberals more than they do in conservatives.

Different patterns were found for the two event-related potentials used in this study, the N400 and the LPP. This was not predicted, but it raises interesting questions for future research. The only N400 pattern found (in average waveforms but not significant in peak minimum comparisons) was in liberal reactions to anti-Care-Fairness statements, suggesting that these statements sounded semantically incongruous to liberals in the manner of “He took his coffee with cream and dog” (Kutas & Hillyard, 1980). Late positive potentials were found for both liberal and conservative reactions to Loyalty-Authority-Sanctity statements, when they rejected the foundation values for conservatives, and when they rejected *or* supported the values for liberals. The fact that LPP, a component associated with affective arousal and attention reallocation in response to negative stimuli, was found in LAS but not CF statements could suggest that group-focused morality may be particularly passion-arousing for both sides of the ideological divide.

Although it gives the first evidence for our hypotheses using neural activity, the results of this study should be taken as initial, not definitive, evidence. Although it was one of the strongest visible ERP patterns on the averaged waveforms, the analyses of minimum peaks did not reveal any difference between critical and control trials to support the liberal N400 reaction to anti-CF sentences. Alternate analysis approaches may also be possible. For instance, Van Berkum et al.

(2009) examined only trials for which participants strongly agreed or disagreed; examining such trials in isolation could tell us more about the time-course of self-reported passionate moral approval or disapproval in line with one or more foundations. Future ERP studies could more closely examine the foundations in isolation, or tradeoffs between the foundations. In general, EEG methods should prove very useful for an in-depth look (with high temporal resolution) at intuitive moral reactions.

### General Discussion

The studies presented here provided evidence that liberals have some intuitive moral reactions that differ from their consciously-endorsed beliefs and ideologies. No such evidence was found for conservatives. These descriptive findings are anticipated by liberal and conservative writings about the normative status of gut-level moral intuitions. For instance, on the right, Leon Kass (1997) famously invoked the “wisdom of repugnance,” advocating listening to one’s intuitive and emotional impulses in deciding right and wrong, concluding “shallow is the soul who has forgotten how to shudder.” On the left, Martha Nussbaum (2010) has urged people to ignore their feelings of disgust when deciding right and wrong: “The politics of disgust is profoundly at odds with the abstract idea of a society based on the equality of all citizens, in which all have a right to the equal protection of the laws. It says that the mere fact that you happen to make me want to vomit is reason enough for me to treat you as a social pariah, denying you some of your most basic rights as a citizen” (p. xiv). While disgust may be most strongly related to the Sanctity foundation, the present studies show evidence of consciously unendorsed “warning bells” related to Loyalty and Authority as well.

The present studies identified these group-focused moral concerns as the locus of liberal “political correction,” which raises the still-open question of whether there may be cases in

which conservatives suppress or correct intuitive reactions to Care and Fairness when they conflict with their ideological commitments. The studies presented here did not find any evidence of this, and suggest that conservatives may be less likely to correct or adjust their affective reactions in moral judgment – a potentially important aspect of ideological differences and culture-war arguments. But future studies may still turn up evidence of conservatives looking or acting like liberals implicitly in a way they do not explicitly.<sup>7</sup>

### *Future Directions*

Although these studies provide multi-method evidence about implicit morality, it is clear that they represent a first step. Moral philosophers (Huemer, 2005; Prinz, 2007) and psychologists (Haidt, 2001; Hoffman, 1987) have become increasingly interested in the role of intuition in moral judgment, and political psychologists have begun examining nonconscious aspects of ideological attitudes (Jost et al., 2008; Nosek et al., 2010). Both of these developments are likely to continue to transform psychological conceptions of morality and ideology as researchers make more use of the measurement tools of implicit social cognition (Hassin et al., 2005). Such work could address open questions regarding what precisely the “warning bell” in liberals might be, for different moral concerns and in different social situations (e.g., is it a flash of disgust, an implicit association, an acquired aversion, an innately-prepared moral intuition, etc.). The present studies show evidence of reactions to morally-relevant stimuli on three different implicit measures, but it’s possible that some intuitive reactions to, say, Sanctity violations involve different physiological and cognitive processes than do reactions to other kinds of moral violations.

---

<sup>7</sup> Wright and Baril (2011) found that conservatives under cognitive load or ego depletion showed reduced concerns about LAS on the MFQ, hence making them “look” more like liberals. However, two studies (Ns 153 and 2396) using two different samples have failed to replicate this finding (Graham, Haidt, & Nosek, 2012).

An important limitation of these studies is their reliance on self-reported placement on a liberal-conservative continuum as the measure of ideology. This brings up a critical question for the present studies: Why rely on self-report for ideology when attempting to bypass self-report for morality variables? First, implicit political identification (as measured by a liberal/conservative self/other IAT) tends to correlate fairly highly with self-reported political identification ( $r \sim .6$ ; see e.g. Graham et al., 2009, Study 2). Second, the present studies focus on explaining explicit, deliberately argued moral debates in the culture war, searching for the implicit moral roots of elaborated political arguments; thus, the primary theoretical interest is in how implicitly-measured moral impulses might vary with self-reported, not implicitly-measured, political ideology. Nevertheless, implicit measurement of ideological identification may be an important next step in understanding the implicit and explicit interactions of morality and ideology (see Nosek et al., 2010).

Not everyone is willing to place themselves on a liberal-conservative continuum, especially when there are other alternative answers to the political orientation question. In fact, this is shown in the above studies: when the ideology item only contained the 7-point liberal-conservative spectrum, the vast majority placed themselves on it (97.5% in Study 2, 99.1% in Study 3). However, when other options were available – don't know/not political, libertarian, and other – only 81.2% placed themselves onto the liberal-conservative continuum (Study 1). Future work applying implicit social cognition methods to the study of ideology should develop new ways of measuring ideological commitments without relying on self-report; in addition, more work should be done on libertarians and others not falling neatly onto a single liberal-conservative dimension.

As a theory of the bases of moral intuitions, Moral Foundations Theory is well-suited to provide theoretical grounding for much of this research. For instance, work is underway examining how foundation-related information about novel individuals (e.g., actions that uphold or violate foundational concerns) can impact implicit attitude formation (Ratliff & Graham, 2012) and involuntary facial movements as measured by electromyography (EMG; Cannon & Schnall, 2011). As Study 4 illustrates, social neuroscience will also provide tools for investigating the nonconscious processes involved in moral judgment; for example, Young, Chakroff, Dungan, Koster-Hale, and Saxe (2012) used functional magnetic-resonance imaging to show differential activation of the right temporo-parietal junction (an area implicated in intentionality assessments) in moral judgments about Care vs. Sanctity. Of course, the topic of implicit morality is large enough to encompass many different theoretical perspectives, not just Moral Foundations Theory.

### *Conclusion*

These studies provide evidence of moral concerns about Loyalty/betrayal, Authority/subversion, and Sanctity/degradation in liberals that are not captured by explicit measures. While ideological differences can be found in different brainwave patterns just milliseconds after exposure to a morally-relevant word, assessing intuitive moral reactions can also uncover similarities between liberals and conservatives not suggested by their consciously-endorsed moral opinions. Understanding precisely which moral intuitions unite ideological opponents, and the timepoints and contexts in which those intuitions go in different directions, could not only aid in understanding the nature of ideological differences, but could aid in negotiating more fruitful cooperation between those with different explicit moralities as well.

## References

- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*, 323-370.
- Cacioppo, J. T., Crites, S. L. Jr., Gardner, W. L., & Berntson, G. G. (1994). Bioelectrical echoes from evaluative categorizations: I. A late positive brain potential that varies as a function of trait negativity and extremity. *Journal of Personality and Social Psychology, 67*, 115-125.
- Cannon, P. & Schnall, S. (2010). Transgressions and expressions: Affective facial muscle activity predicts moral judgments. Manuscript submitted for publication.
- Delorme, A. & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods, 134*, 9-21.
- Eidelman, S., Crandall, C.S., Goodman, J.A., & Blanchar, J.C. (in press). Low-effort thought promotes political conservatism. *Personality and Social Psychology Bulletin*.
- Ferguson, M. J. (2007). On the automatic evaluation of end-states. *Journal of Personality and Social Psychology, 92*, 596-611.
- Fiske, A. P. (1992). Four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review, 99*, 689-723.
- Gawronski, B., & Payne, B. K., Eds. (2010). *Handbook of implicit social cognition: Measurement, theory, and applications*. New York: Guilford Press.
- Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.
- Graham, J., & Haidt, J. (2010). Beyond beliefs: Religions bind individuals into moral communities. *Personality and Social Psychology Review, 14*, 140-150

- Graham, J., & Haidt, J. (in press). Sacred values and evil adversaries: A moral foundations approach. in P. Shaver & M. Mikulincer (Eds.), *The Social Psychology of Morality: Exploring the Causes of Good and Evil*. New York: APA Books.
- Graham, J., Haidt, J. & Nosek, B.A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*, 1029-1046.
- Graham, J., Haidt, J. & Nosek, B.A. (2012). *Cognitive load effects on Moral Foundations Questionnaire scores*. Unpublished data, University of Southern California.
- Graham, J., Haidt, J., & Rimm-Kaufman, S. E. (2008). Ideology and intuition in moral education. *European Journal of Developmental Science*, *2*, 269-286.
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, K., & Ditto, P. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, *101*, 366-385.
- Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., & Cohen, J.D. (2001). An fMRI investigation of emotional engagement in moral Judgment. *Science*, *293*, 2105-2108.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4-27.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*, 814-834.
- Haidt, J. (2012). *The Righteous Mind: Why Good People are Divided by Politics and Religion*. New York: Pantheon/Knopf.
- Haidt, J. & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, *20*, 98-116.
- Haidt, J. & Graham, J. (2009). The planet of the Durkheimians, where community, authority and sacredness are foundations of morality. In J. T. Jost, A. C. Kay, & H. Thorisdottir (Eds.),

- Social and Psychological Bases of Ideology and System Justification* (pp. 371-401). New York: Oxford University Press.
- Haidt, J., Graham, J., & Joseph, C. (2009). Above and below left-right: Ideological narratives and moral foundations. *Psychological Inquiry, 20*, 110-119.
- Haidt, J., & Hersh, M. (2001). Sexual morality: The cultures and emotions of conservatives and liberals. *Journal of Applied Social Psychology, 31*, 191-221.
- Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus, 55-66*.
- Hassin, R. R., Uleman, J. S., & Bargh, J. A. (Eds.) (2005). *The New Unconscious*. New York: Oxford University Press.
- Hoffman, M. L. (1987). The contribution of empathy to justice and moral judgment. In N. Eisenberg and J. Strayer (Eds.), *Empathy and its development* (pp. 47-80). New York: Cambridge University Press.
- Huemer, M. (2005). *Ethical Intuitionism*. Hampshire, UK: Palgrave Macmillan.
- Hunter, J. D. (1991). *Culture Wars: The Struggle to Define America*. New York: Basic Books.
- Inbar, Y., Pizarro, D.A., Knobe, J., & Bloom, P. (2009). Disgust sensitivity predicts intuitive disapproval of gays. *Emotion, 9*, 435-439.
- Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology, 75*, 887-900.
- Jost, J. T. (2006). The end of the end of ideology. *American Psychologist, 61*, 651-670.
- Jost, J.T., Federico, C.M. & Napier, J.L. (2009). Political ideology: Its structure, functions, and elective affinities. *Annual Review of Psychology, 60*, 307-333.

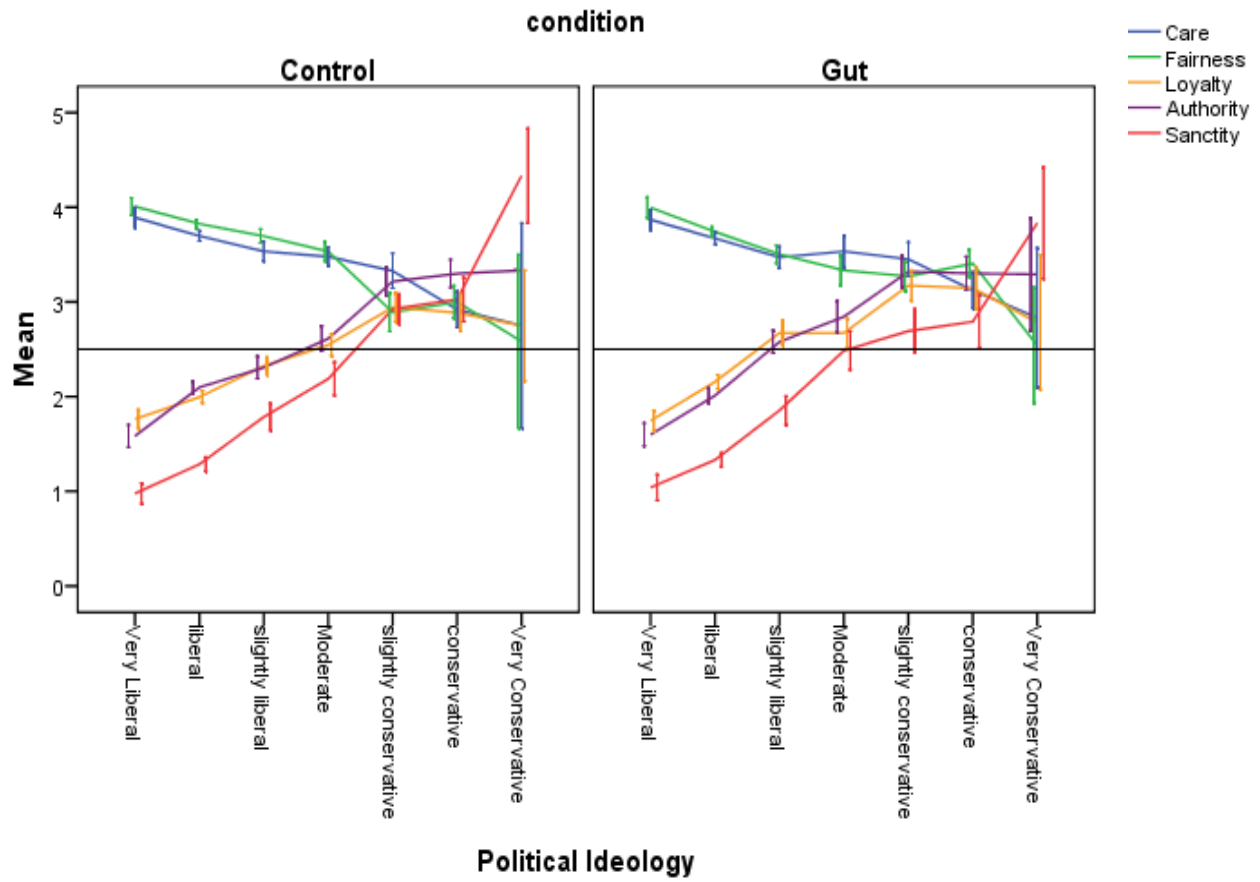
- Jost, J. T., Kay, A. C., & Thorisdottir, H. (Eds.). (2009). *Social and Psychological Bases of Ideology and System Justification*. New York, NY: Oxford University Press.
- Jost, J. T., Nosek, B.A. & Gosling, S.D. (2008). Ideology: Its resurgence in social, personality, and political psychology. *Perspectives on Psychological Science*, 3, 126-136.
- Kass, L. R. (1997). The wisdom of repugnance. *The New Republic*, 216, 17-26.
- Kutas, M., & Hillyard, S.A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207, 203-205.
- Kutas, M., & Hillyard, S.A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307, 161-163.
- Luo, Q., Nakic, M., Wheatley, T., Richell, R., Martin, A., & Blair, R.J.R. (2006) The neural basis of implicit moral attitude: an IAT study using event-related fMRI. *NeuroImage*, 30, 1449-57.
- McAdams, D. P., Albaugh, M., Farber, E., Daniels, J., Logan, R. L., & Olson, B. (2008). Family metaphors and moral intuitions: How conservatives and liberals narrate their lives. *Journal of Personality and Social Psychology*, 95, 978-990.
- Morris, J. P., Squires, N. K., Taber, C. S., & Lodge, M. (2003). Activation of political attitudes: A psychophysiological examination of the hot cognition hypothesis. *Political Psychology*, 24, 727-745.
- Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology*, 77, 167-184.

- Nosek, B. A., Graham, J., & Hawkins, C. B. (2010). Implicit political cognition. In B. Gawronski & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications* (pp. 548-564). New York: Guilford.
- Nosek, B.A., Hawkins, C.B., & Frazier, R.S. (2011). Implicit social cognition: Methods and mechanisms. *Trends in Cognitive Sciences, 15*, 152-159.
- Nussbaum, M. (2010). *From Disgust to Humanity: Sexual Orientation and Constitutional Law*. New York: Oxford University Press.
- Oppenheimer, D., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45*, 867-872.
- Payne, B.K., Cheng, C. M., Govorun, O., & Stewart, B. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89*, 277-293.
- Prinz, J. (2007). *The emotional construction of morals*. Oxford, UK: Oxford University Press.
- Ratliff, K. A., & Graham, J. (2012). Implicit attitude formation based on moral information. Manuscript in preparation.
- Ranganath, K. A., Smith, C. T., & Nosek, B. A. (2008). Distinguishing automatic and controlled components of attitudes from indirect and direct measurement methods. *Journal of Experimental Social Psychology, 44*, 386-396.
- Rozin, P., & Royzman, E. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review, 5*, 296-320.
- Schnall, S., Benton, J., & Harvey, S. (2008). With a clean conscience: Cleanliness reduces the severity of moral judgments. *Psychological Science, 19*, 1219-1222.

- Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin, 34*, 1096-1109.
- Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). The "big three" of morality (autonomy, community, and divinity), and the "big three" explanations of suffering. In A. Brandt, & P. Rozin (Eds.), *Morality and health* (pp. 119-169). New York: Routledge.
- Skitka, L. J., & Tetlock, P. E. (1993). Providing public assistance: Cognitive and motivational processes underlying liberal and conservative policy preferences. *Journal of Personality and Social Psychology, 65*, 1205-1223.
- Skitka, L. J., Mullen, E., Griffin, T., Hutchinson, S., & Chamberlin, B. (2002). Dispositions, scripts, or motivated correction? understanding ideological differences in explanations for social problems. *Journal of Personality and Social Psychology, 83*, 470-487.
- Smith, C., & Vaisey, S. (2010). Moral foundations and charitable giving in a nationally-representative sample. Manuscript in preparation.
- Van Berkum, J.J.A. (2008). Understanding sentences in context: What brain waves can tell us. *Current Directions in Psychological Science, 17*, 376-380.
- Van Berkum, J. J. A., Holleman, B., Nieuwland, M., Otten, M., & Murre, J. (2009). Right or wrong? The brain's fast response to morally objectionable statements. *Psychological Science, 20*, 1092 -1099.
- Van Leeuwen, F., & Park, J. H. (2009). Perceptions of social dangers, moral foundations, and political orientation. *Personality and Individual Differences, 47*, 169-173.
- Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science, 16*, 780-784.

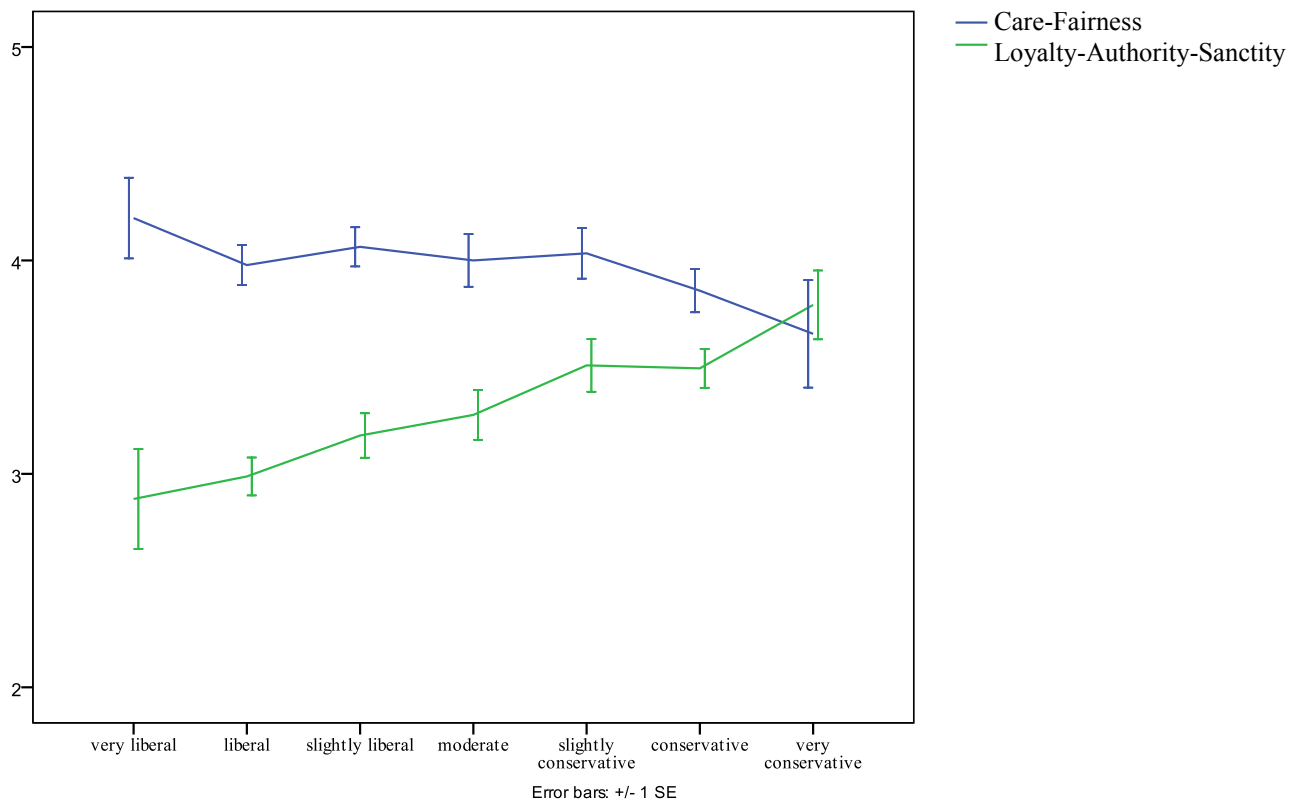
- Wilson, T. D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious*. Cambridge, MA: Belknap Press.
- Wright, J.C., & Baril, G. (2011). The role of cognitive resources in determining our moral intuitions: Are we all liberals at heart? *Journal of Experimental Social Psychology, 47*, 1007-1012.
- Young, L., Camprodon, J., 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. *PNAS, 107*, 6753-6758.
- Young, L., Chakroff, A., Dungan, J., Koster-Hale, J., Saxe, R. (2012). When are we mindful of moral minds? A different role for theory of mind in evaluating assault versus incest. Manuscript under review.
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*, 151-175.
- Zhong, C. B., Strejcek, B., & Sivanathan, N. (2010). A clean self can render harsh moral judgment. *Journal of Experimental Social Psychology, 46*, 859-862.

Figure 1. Moral Foundations endorsements across political ideology, Study 1



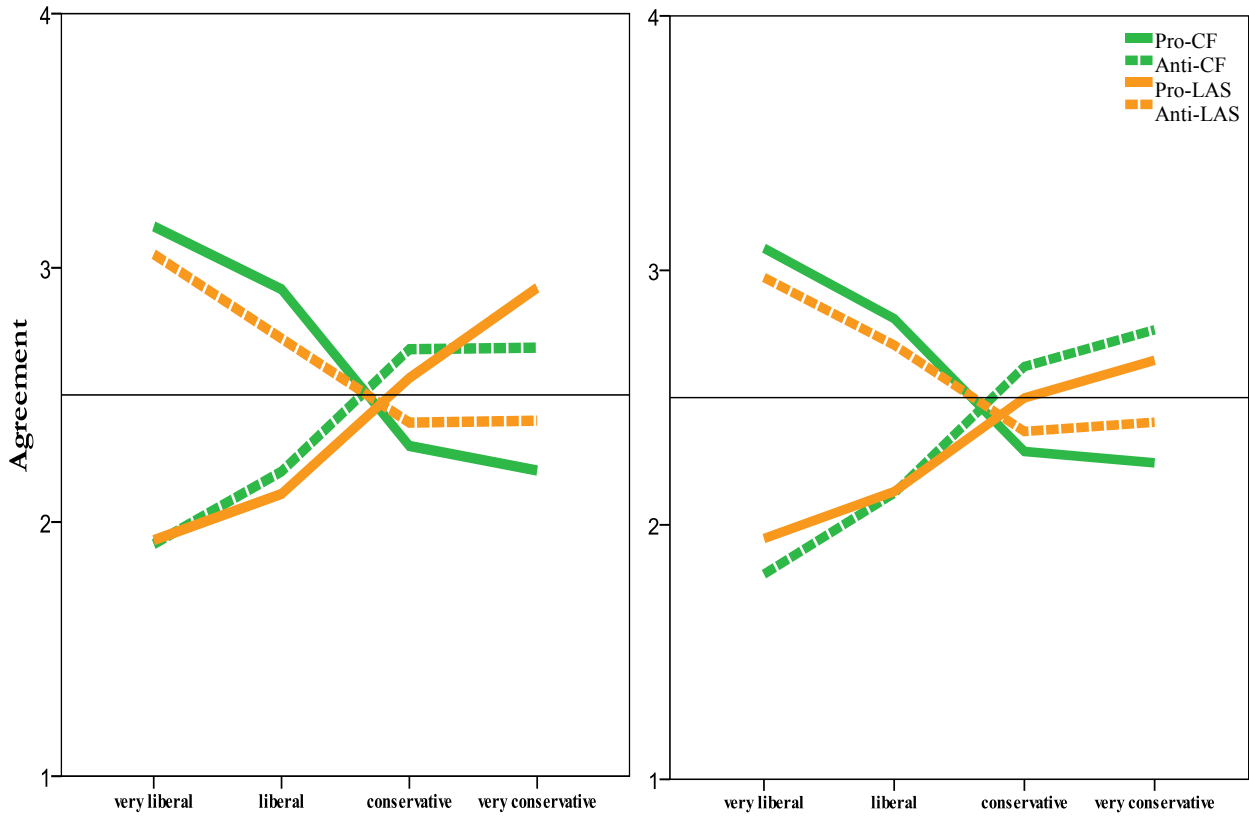
Note. Error bars represent standard error of the mean.

Figure 2. Ratings of foundation-related words across political ideology, Study 2



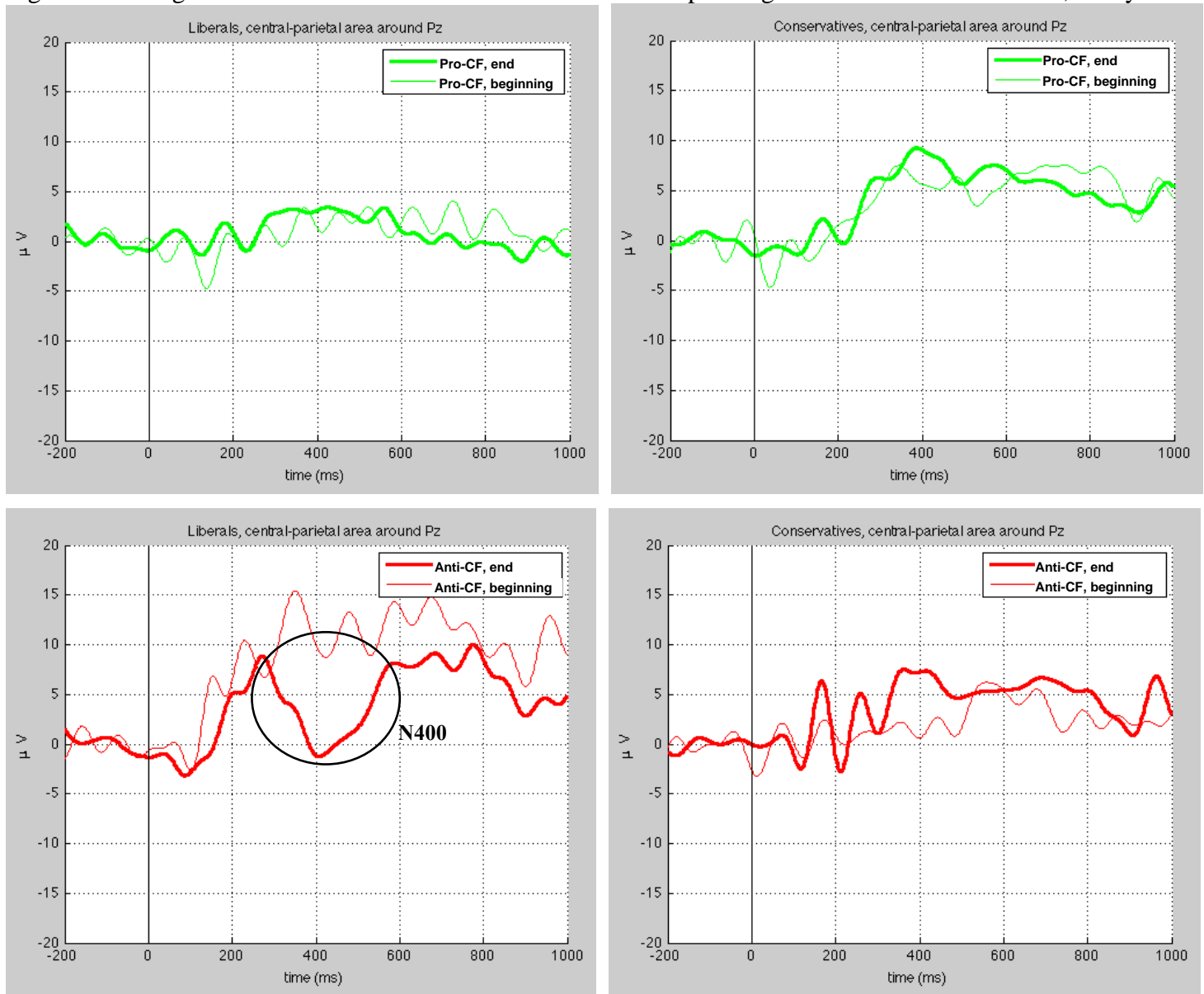
*Note.* Explicit indices were calculated by averaging ratings of all the vice words of the relevant foundations and subtracting it from averaged ratings of all the virtue words. Higher numbers indicate greater preference for foundation-related virtues over foundation-related vices. All words were rated on a 1-7 scale, and can be found in Supplemental Appendix A.

Figure 3. Agreement and disagreement with moral sentences, Study 4



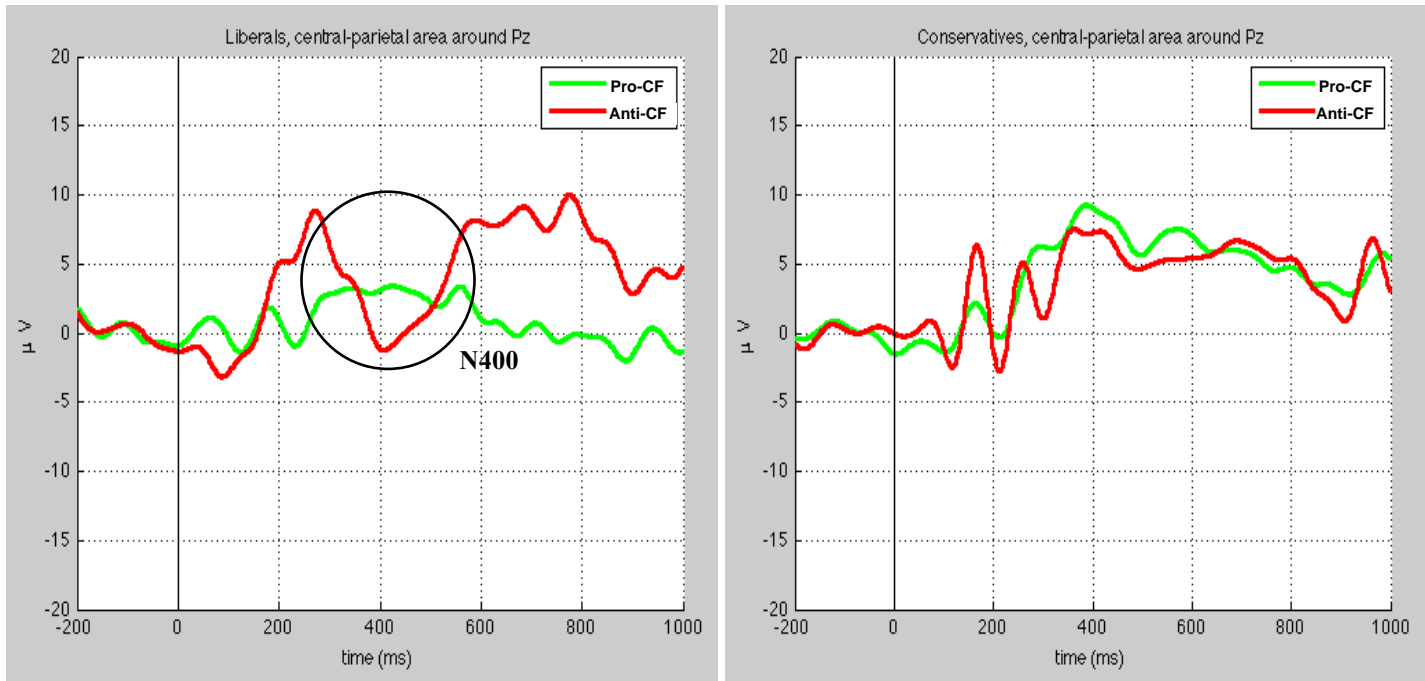
*Note.* Line at 2.5 represents agreement/disagreement border; 2=Slightly disagree and 3=Slightly agree. Left panel shows sentences in which critical words came near the end of the sentence, right panel shows sentences in which critical words came near the beginning of the sentence.

Figure 4. Averaged waveforms for liberals and conservatives responding to Care-Fairness sentences, Study 4



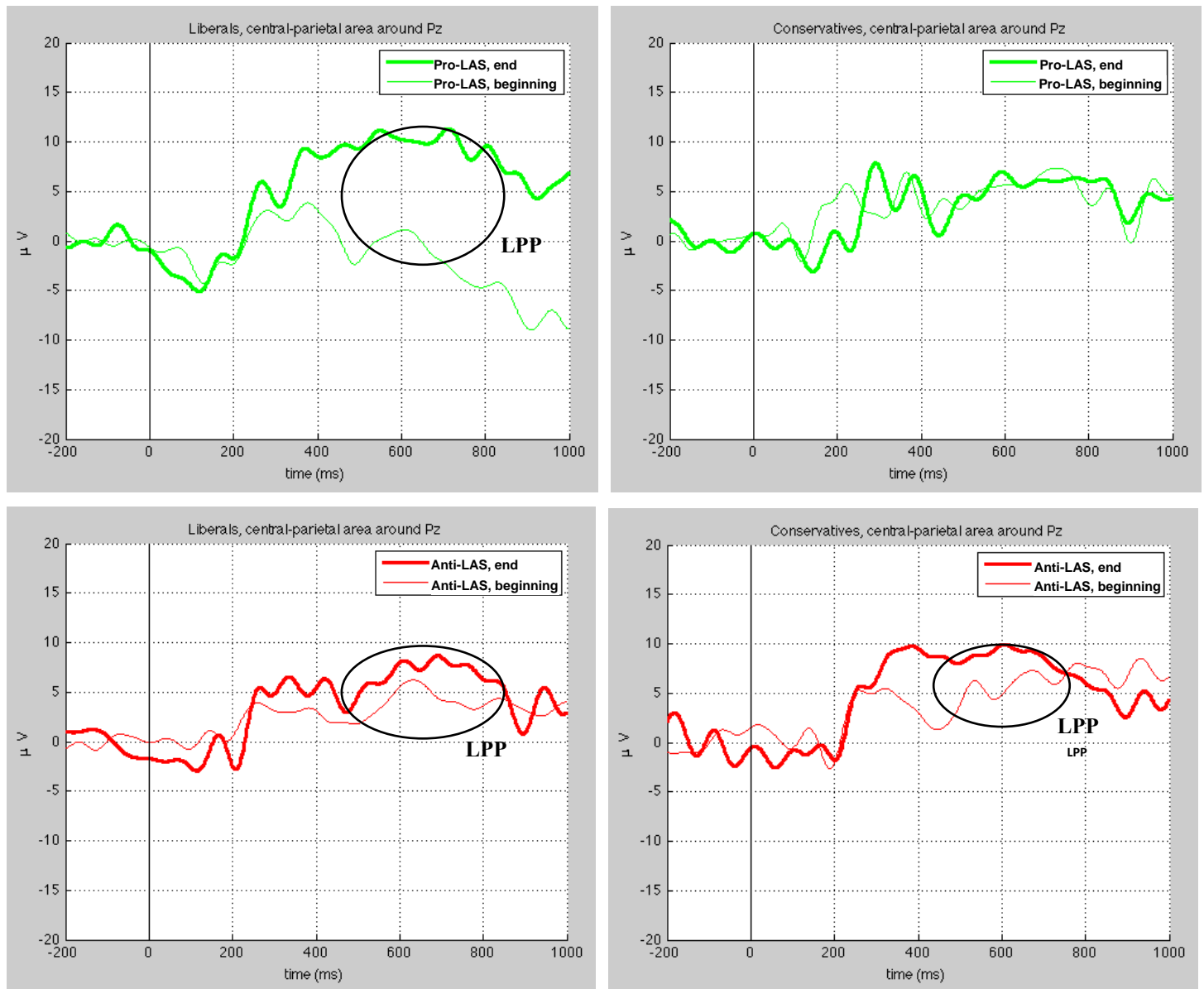
*Note.* Green lines denote critical words in pro-CF sentences, red lines denote critical words in anti-CF sentences. Bold lines indicate when critical words appeared at the end of the sentence (where moral and semantic meaning of their endorsement/nonendorsement was clear), and thin lines indicate when the same critical words appeared at the beginning of the sentence, before their meaning was apparent. N400 negative potential (liberal reaction to anti-CF endorsements at the end of the word, but not the beginning) is highlighted in bottom left panel.

Figure 5. Comparison of reactions to Pro-CF and Anti-CF endorsements, Study 4



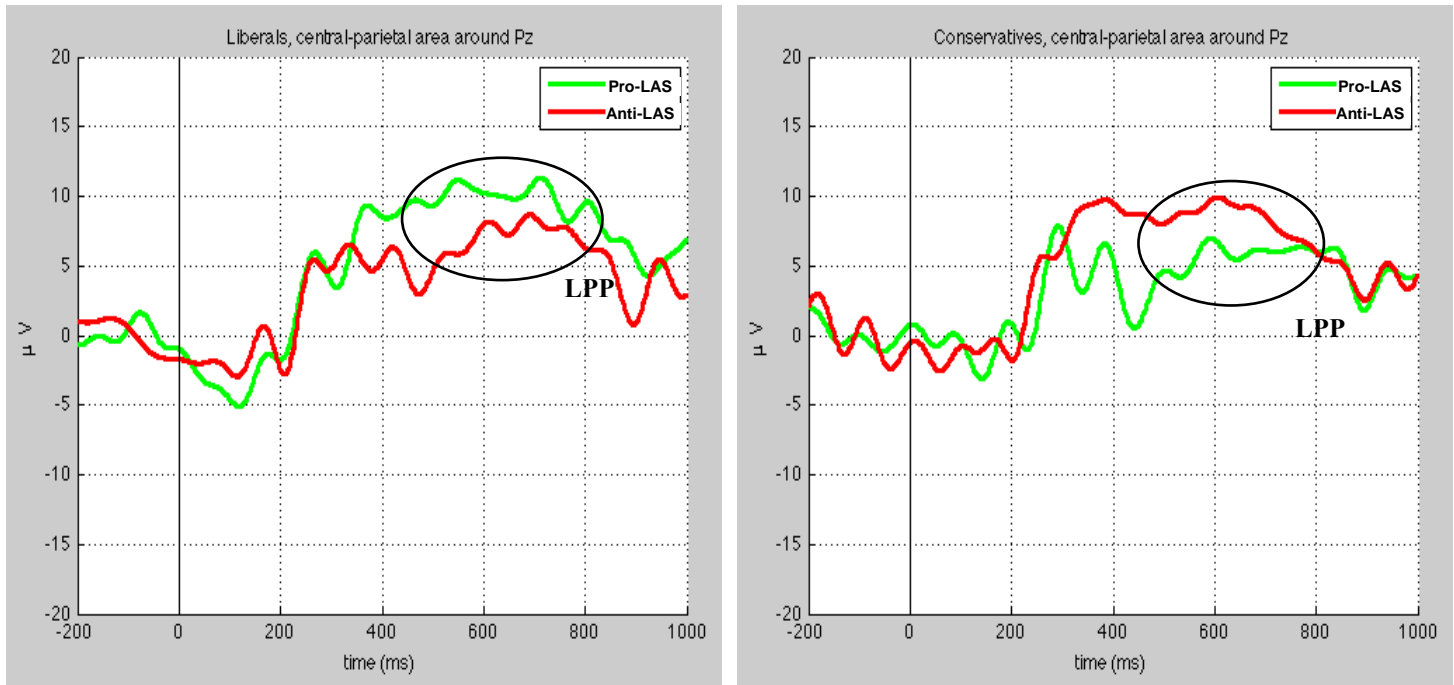
*Note.* Green lines denote critical words in pro-CF sentences, red lines denote critical words in anti-CF sentences. Critical words appeared at the end of both types of sentences. N400 negative potential (liberal reaction to anti-CF endorsements) is highlighted in left panel.

Figure 6. Averaged waveforms for liberals and conservatives responding to Loyalty-Authority-Sanctity sentences, Study 4



*Note.* Green lines denote critical words in pro-LAS sentences, red lines denote critical words in anti-LAS sentences. Bold lines indicate when critical words appeared at the end of the sentence (where moral and semantic meaning of their endorsement/nonendorsement was clear), and thin lines indicate when the same critical words appeared at the beginning of the sentence, before their meaning was apparent. N400 negative potential (liberal reaction to anti-CF endorsements at the end of the word, but not the beginning) is highlighted in bottom left panel.

Figure 7. Comparison of reactions to Pro-LAS and Anti-LAS endorsements, Study 4



*Note.* Green lines denote critical words in pro-LAS sentences, red lines denote critical words in anti-LAS sentences. Critical words appeared at the end of both types of sentences. Late positive potentials (LPP) are shown in both panels. Note that for liberals the stronger reaction is to Pro-LAS sentences, and for conservatives the stronger reaction is to Anti-LAS sentences.

**Supplemental Appendix A. Foundation-related words used in Study 2**

Care-Fairness primes:	hurt, cruel, cheat,
Loyalty-Authority-Sanctity primes:	traitor, revolt, sin
Control primes:	table, chair, running, listen, watch, fruit, green, sofa, city, shelf, grocery, swimming, blender
Positive target adjectives:	excellent, wonderful, terrific, great, super, happy, nice, fantastic, polite, pleasant, lovely, delightful
Negative target adjectives:	horrible, awful, dangerous, sad, dreadful, terrible, atrocious, unpleasant, poor, dire, unhappy, grim
Explicitly-rated words:	
Care-virtue:	nurturing, care, protect, compassionate
Care-vice:	cruel, hurt, aggressive, war
Fairness-virtue:	equality, justice, rights, fair
Fairness-vice:	cheat, bias, prejudice, unfair
Loyalty-virtue:	community, patriotic, united, loyal
Loyalty-vice:	foreign, traitor, treason, betray
Authority-virtue:	respect, authority, duty, obey
Authority-vice:	anarchy, dissent, revolt, subversive
Sanctity-virtue:	pure, wholesome, chastity, sacred
Sanctity-vice:	sin, slut, disgusting, dirty

Supplemental Appendix B. Pictures used in Affect Misattribution Procedure, Study 3

Care virtues



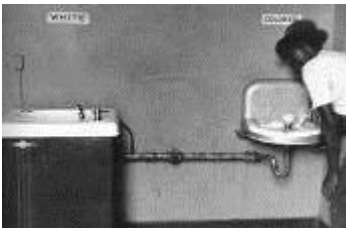
Care Vices



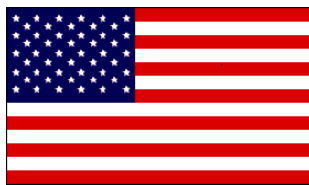
Fairness Virtues



Fairness Vices



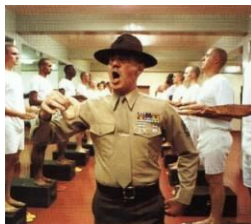
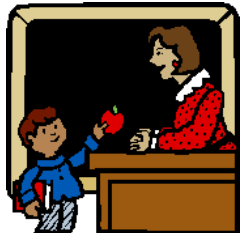
Loyalty Virtues



Loyalty Vices



Authority Virtues



Authority Vices



Sanctity Virtues



Sanctity Vices



### Supplemental Appendix C. Sentences used in EEG Paradigm, Study 4

[First sentence in each block of four shows the critical word (e.g., “necessary” in sentence 1) at or near the end of the sentence. Each participant saw either the first and third (or second and fourth) sentence in each block.]

#### Sentences supporting/rejecting Care-Fairness:

1. Showing compassion to homeless drug addicts is necessary.
2. It is necessary to show compassion to homeless drug addicts.
3. Showing compassion to homeless drug addicts is futile.
4. It is futile to show compassion to homeless drug addicts.
  
5. Protecting one's home with lethal force is unacceptable.
6. It is unacceptable to protect one's home with lethal force.
7. Protecting one's home with lethal force is acceptable.
8. It is acceptable to protect one's home with lethal force.
  
9. Hurting other people’s feelings is inexcusable.
10. It is inexcusable to hurt other people's feelings.
11. Hurting other people's feelings can be necessary.
12. It is necessary to hurt other people’s feelings.
  
13. Inflicting excruciating physical harm in self-defense is unwarranted.
14. It is unwarranted to inflict excruciating physical harm in self-defense.
15. Inflicting excruciating physical harm in self-defense is warranted.
16. It is warranted to inflict excruciating physical harm in self-defense.
  
17. People who lobby for animal rights are admirable.
18. Admirable people lobby for animal rights.
19. People who lobby for animal rights are naïve.
20. Naïve people lobby for animal rights.
  
21. Keeping a firearm in your house for self-defense is risky.
22. It is risky to keep a firearm in your house for self-defense.
23. Keeping a firearm in your house for self-defense is wise.
24. It is wise to keep a firearm in your house for self-defense.
  
25. Attempting to rescue stray or homeless dogs is noble.
26. It is noble to attempt to rescue stray or homeless dogs.
27. Attempting to rescue stray or homeless dogs is pointless.
28. It is pointless to attempt to rescue stray or homeless dogs.
  
29. Spanking children who misbehave is cruel.
30. It is cruel to spank children who misbehave.
31. Spanking children who misbehave is necessary.
32. It is necessary to spank children who misbehave.
  
33. Placing more restrictions on the hunting of animals is necessary.
34. It is necessary to place more restrictions on the hunting of animals.
35. Placing more restrictions on the hunting of animals is unnecessary.

36. It is unnecessary to place more restrictions on the hunting of animals.
37. Slapping a child for misbehaving is wrong.
38. It is wrong to slap a child for misbehaving.
39. Slapping a child for misbehaving is acceptable.
40. It is acceptable to slap a child for misbehaving.
41. War can be justified never.
42. Never can war be justified.
43. War can be justified sometimes.
44. Sometimes war can be justified.
45. The government's efforts to protect vulnerable members of society are insufficient.
46. The government's efforts are insufficient in protecting vulnerable members of society.
47. The government's efforts to protect vulnerable members of society are sufficient.
48. The government's efforts are sufficient in protecting vulnerable members of society.
49. Capital punishment is morally wrong.
50. It is morally wrong to allow capital punishment.
51. Capital punishment is morally justified.
52. It is morally justified to allow capital punishment.
53. The goal of equal wealth in society is noble
54. It is noble to have the goal of equal wealth in society
55. The goal of equal wealth in society is misguided
56. It is misguided to have the goal of equal wealth in society
57. Higher taxes on the rich are fair
58. It is fair to place higher taxes on the rich
59. Higher taxes on the rich are unfair
60. It is unfair to place higher taxes on the rich
61. Affirmative action for women and minorities is necessary
62. It is necessary to have affirmative action for women and minorities
63. Affirmative action for women and minorities is wrong
64. It is wrong to have affirmative action for women and minorities
65. Making sure that all citizens have health insurance is necessary
66. It is necessary to make sure that all citizens have health insurance
67. Making sure that all citizens have health insurance is unnecessary
68. It is unnecessary to make sure that all citizens have health insurance
69. People accused of a crime must be treated more fairly
70. We must treat more fairly people accused of a crime
71. People accused of a crime must be treated more firmly
72. We must treat more firmly people accused of a crime
73. Giving preference to college applicants from disadvantaged school districts is desirable
74. It is desirable to give preference to college applicants from disadvantaged school districts
75. Giving preference to college applicants from disadvantaged school districts is objectionable

76. It is objectionable to give preference to college applicants from disadvantaged school districts
  77. Total equality in the workplace is necessary
  78. It is necessary to have total equality in the workplace
  79. Total equality in the workplace is unrealistic
  80. It is unrealistic to have total equality in the workplace
  81. Affirmative action in college admissions is fair
  82. It is fair to have affirmative action in college admissions
  83. Affirmative action in college admissions is unfair
  84. It is unfair to have affirmative action in college admissions
  85. Today feminism is important
  86. An important cause today is feminism
  87. Today feminism is unimportant
  88. An unimportant cause today is feminism
  89. Income redistribution from the rich to the poor is fair
  90. It is fair to redistribute income from the rich to the poor
  91. Income redistribution from the rich to the poor is unfair
  92. It is unfair to redistribute income from the rich to the poor
  93. Curtailing individual rights in the name of national security is wrong
  94. It is wrong to curtail individual rights in the name of national security
  95. Curtailing individual rights in the name of national security is necessary
  96. It is necessary to curtail individual rights in the name of national security
  97. People who fight for social justice and equality are brave
  98. Brave people fight for social justice and equality
  99. People who fight for social justice and equality are annoying
  100. Annoying people fight for social justice and equality
- Sentences supporting/rejecting Loyalty-Authority-Sanctity:
101. Someone who lacks patriotism is wrong
  102. It is wrong for someone to lack patriotism
  103. Someone who lacks patriotism is fine
  104. It is fine for someone to lack patriotism
  105. Someone who lacks school spirit is wrong
  106. It is wrong for someone to lack school spirit
  107. Someone who lacks school spirit is fine
  108. It is fine for someone to lack school spirit
  109. When I see UVA colors it makes me feel proud
  110. It makes me feel proud when I see UVA colors
  111. When I see UVA colors it makes me feel annoyed
  112. It makes me feel annoyed when I see UVA colors
  113. When other people point out my country's flaws, I usually disagree

114. I usually disagree when other people point out my country's flaws
115. When other people point out my country's flaws, I usually agree
116. I usually agree when other people point out my country's flaws
  
117. It should be illegal to burn the American flag
118. Burning the American flag should be illegal
119. It should be legal to burn the American flag
120. Burning the American flag should be legal
  
121. Protesting the actions of your government is disrespectful
122. It is disrespectful to protest the actions of your government
123. Protesting the actions of your government is admirable
124. It is admirable to protest the actions of your government
  
125. Protesting the actions of your company is disrespectful
126. It is disrespectful to protest the actions of your company
127. Protesting the actions of your company is admirable
128. It is admirable to protest the actions of your company
  
129. Traitors to our nation should be punished
130. We should punish traitors to our nation
131. Traitors to our nation should be forgiven
132. We should forgive traitors to our nation
  
133. Abstaining from saying the Pledge of Allegiance is unacceptable
134. It is unacceptable to abstain from saying the Pledge of Allegiance
135. Abstaining from saying the Pledge of Allegiance is acceptable
136. It is acceptable to abstain from saying the Pledge of Allegiance
  
137. Dissent is traitorous
138. It is traitorous to dissent
139. Dissent is patriotic
140. It is patriotic to dissent
  
141. When I hear foreigners say they hate America, I find it infuriating
142. I find it infuriating when I hear foreigners say they hate America
143. When I hear foreigners say they hate America, I find it understandable
144. I find it understandable when I hear foreigners say they hate America
  
145. Allowing more immigrants into America will hurt us in the long run
146. It will hurt us in the long run to allow more immigrants into America
147. Allowing more immigrants into America will benefit us in the long run
148. It will benefit us in the long run to allow more immigrants into America
  
149. Figures of authority should be obeyed
150. We should obey figures of authority
151. Figures of authority should be questioned
152. We should question figures of authority

153. Listening to authority figures, we should be trusting
154. We should trust authority figures
155. Listening to authority figures, we should be skeptical
156. We should be skeptical of authority figures
  
157. Criminals should be punished
158. Punishment should be the treatment for criminals
159. Criminals should be rehabilitated
160. Rehabilitation should be the treatment for criminals
  
161. Questioning an order from a superior is immoral
162. It is immoral to question an order from a superior
163. Questioning an order from a superior is moral
164. It is moral to question an order from a superior
  
165. Disobeying the directives of a coach or group leader is unacceptable
166. It is unacceptable to disobey the directives of a coach or group leader
167. Disobeying the directives of a coach or group leader is acceptable
168. It is acceptable to disobey the directives of a coach or group leader
  
169. In the teenage years, parental advice should be heeded
170. Teenagers should heed their parents' advice
171. In the teenage years, parental advice should be questioned
172. Teenagers should question their parents' advice
  
173. If you feel that a law is unjust, you should follow it
174. You should follow a law you feel is unjust
175. If you feel that a law is unjust, you should break it
176. You should break a law you feel is unjust
  
177. Correcting a teacher if he or she makes a mistake is disrespectful
178. It is disrespectful to correct a teacher if he or she makes a mistake
179. Correcting a teacher if he or she makes a mistake is good
180. It is good to correct a teacher if he or she makes a mistake
  
181. Talking back to one's parents is immoral
182. It is immoral to talk back to one's parents
183. Talking back to one's parents is necessary
184. It is necessary to talk back to one's parents
  
185. Always doing as you're told is moral
186. It is moral to always do as you're told
187. Always doing as you're told is immoral
188. It is immoral to always do as you're told
  
189. Fighting for political revolution is misguided
190. It is misguided to fight for political revolution
191. Fighting for political revolution is necessary
192. It is necessary to fight for political revolution

- 193. When a protest becomes too noisy, police action to restrain the crowd is wrong
- 194. It is wrong for police to take action to restrain the crowd when a protest becomes too noisy
- 195. When a protest becomes too noisy, police action to restrain the crowd is justified
- 196. It is justified for police to take action to restrain the crowd when a protest becomes too noisy

- 197. Parents who let their teenagers explore and experiment are misguided
- 198. Parents are misguided to allow teenage children to explore and experiment
- 199. Parents who let their teenagers explore and experiment are wise
- 200. Parents are wise to allow teenage children to explore and experiment

- 201. Fulfilling the traditional duties of your gender is noble
- 202. It is noble to fulfill the traditional duties of your gender
- 203. Fulfilling the traditional duties of your gender is outdated
- 204. It is outdated to fulfill the traditional duties of your gender

- 205. Treating one's body as a temple, not a playground, is necessary
- 206. It is necessary to treat one's body as a temple, not a playground
- 207. Treating one's body as a temple, not a playground, is outdated
- 208. It is outdated to treat one's body as a temple, not a playground

- 209. Tattooing or piercing one's body is disgusting
- 210. It is disgusting to tattoo or pierce one's body
- 211. Tattooing or piercing one's body is exciting
- 212. It is exciting to tattoo or pierce one's body

- 213. Casual sex is wrong
- 214. It is wrong to engage in casual sex
- 215. Casual sex is fun
- 216. It is fun to engage in casual sex

- 217. Streaking or being nude in public is wrong
- 218. It is wrong to streak or be nude in public
- 219. Streaking or being nude in public is fine
- 220. It is fine to streak or be nude in public

- 221. People engaging in homosexual acts is unnatural
- 222. It is unnatural for people to engage in homosexual acts
- 223. People engaging in homosexual acts is natural
- 224. It is natural for people to engage in homosexual acts

- 225. Permanently altering one's body for decorative purposes is unacceptable
- 226. It is unacceptable to permanently alter one's body for decorative purposes
- 227. Permanently altering one's body for decorative purposes is fine
- 228. It is fine to permanently alter one's body for decorative purposes

- 229. Openly discussing sexuality is improper
- 230. It is improper to openly discuss sexuality
- 231. Openly discussing sexuality is okay

232. It is okay to openly discuss sexuality
233. Recreational use of drugs is unjustifiable
234. It is unjustifiable to recreationally use drugs
235. Recreational use of drugs is justifiable
236. It is justifiable to recreationally use drugs
237. Gambling on a regular basis is unacceptable
238. It is unacceptable to gamble on a regular basis
239. Gambling on a regular basis is acceptable
240. It is acceptable to gamble on a regular basis
241. Banning gay marriage is necessary
242. It is necessary to ban gay marriage
243. Banning gay marriage is wrong
244. It is wrong to ban gay marriage
245. Requiring young women to be immunized against HPV is immoral
246. It is immoral to require young women to be immunized against HPV
247. Requiring young women to be immunized against HPV is necessary
248. It is necessary to require young women to be immunized against HPV
249. Selling my soul for money would be wrong
250. It would be wrong to sell my soul for money
251. Selling my soul for money would be easy
252. It would be easy to sell my soul for money
253. Teaching middle-schoolers about contraception is wrong
254. It is wrong to teach middle-schoolers about contraception
255. Teaching middle-schoolers about contraception is necessary
256. It is necessary to teach middle-schoolers about contraception
257. Embryonic stem cell research is immoral
258. It is immoral to conduct embryonic stem cell research
259. Embryonic stem cell research is moral
260. It is moral to conduct embryonic stem cell research
261. Masturbating is unacceptable
262. It is unacceptable to masturbate
263. Masturbating is acceptable
264. It is acceptable to masturbate