Linking Issue Stances and Trait Inferences: A Theory of Moral Exemplification

Scott Clifford  
Postdoctoral Fellow  
Duke University  
Social Science Research Institute  
140 Science Drive  
Durham, NC 27708  
scott.clifford@duke.edu

Conditionally accepted for publication at the Journal of Politics

Abstract

Considerable research has demonstrated the importance of perceptions of politicians’ character traits for vote choice. Yet, we know little about the antecedents of trait attributions. Drawing on Moral Foundations Theory, I argue that character traits correspond to particular moral foundations. I introduce a theory of moral exemplification, according to which individuals use their own moral motivations, and the character traits exemplifying these motivations, to interpret the behavior of politicians. The analysis of three separate studies reveals support for the theory. First, individuals’ moral foundations predict the accessibility of corresponding traits, and thus their propensity to be used in evaluation. Second, across two experiments, politicians’ issue stances shape perceptions of their traits. As predicted, however, the type of trait inference made depends on the moral foundation associated with the individual’s issue stance. I conclude with a discussion of how moral exemplification theory provides insight into trait ownership theory and campaign strategy.
Issues have little autonomous effect on election outcomes. Rather, issues are vehicles that some House members choose to convey their qualifications, their sense of identification, and their sense of empathy. It is not the statement of an issue position that wins elections, but the presentation of self by the candidate as he states his issue position. […] A good issue for a candidate is, in this view, one that allows him to present himself as a person in a favorable light (Fenno 1978, 134).

Political campaigns tailor their efforts to portraying their candidate as virtuous, exemplifying character traits such as empathy, integrity, and leadership.1 Detailed interviews with Senate campaign managers reveal that a candidate’s character traits are the primary theme of most campaigns (Kahn and Kenney 1999). Evidence from inside a presidential campaign demonstrates that candidates strategically prime character traits (Druckman, Jacobs, and Ostermeier 2004). It is little wonder that campaigns focus on shaping character trait impressions, as a vast literature has demonstrated the importance of trait perceptions for approval and vote choice (e.g., Funk 1996, 1999; Hayes 2005, 2010; Fridkin and Kenney 2011). Indeed, Johnston et al. (2004, 129) argue that “shifts in trait perceptions were the most important single story of the [2000 presidential] campaign.”

Trait perceptions are also tied into partisan stereotypes. According to trait ownership theory, the parties’ ownership of particular issues (Petrocik 1996) creates ownership of associated character traits (Hayes 2005). For example, Democratic politicians’ concern for social welfare issues creates trait ownership over compassion and empathy, while Republican politicians’ concern for foreign policy and family values creates trait ownership over leadership and integrity. Indeed, campaign advertising frequently “dovetails” issue stances with character traits, under the assumption that they are mutually reinforcing (e.g., Kern 1989, Just et al. 1996, Kahn and Kenney 1999). Supporting these claims, Hayes

1 Replication data and an online appendix will be made available at scottaclifford.com upon publication.
(2005) finds that, over twenty-four years of ANES data, Democratic presidential candidates were viewed as more compassionate and empathetic, while Republican candidates were viewed as stronger leaders and having greater integrity.

Underlying our understanding of campaign strategy and partisan trait stereotypes is an assumption about the causal connection between issue stances and trait perceptions. Yet, while we have some evidence that issue stances cause trait inferences (Peterson 2005; Rahn et al. 1990; Rapoport, Metcalf, and Hartman 1989), “we know very little about how trait and issue inferences are linked” (McGraw 2011, 193). Indeed, existing literature offers diverging theories for why issue stances cause trait inferences and provide little guidance for which issue stances cause which trait perceptions. This shortcoming limits our understanding of partisan stereotypes, politicians’ attempts to shape their public images, and how citizens evaluate these images.

In this paper, I build on recent research demonstrating that five moral foundations strongly predict ideological and political attitudes (e.g., Koleva et al. 2012; Weber and Federico 2012). Extending psychological research on Moral Foundations Theory, I introduce a theory of moral exemplification, which holds that each moral foundation is represented by specific character traits. I argue moral foundations shape not only the positions we take on political issues, but also the character traits we perceive to motivate others’ positions. Just as individuals use different moral foundations to assess an issue, they use different character traits to evaluate a politician’s stance on that issue. Across three studies, my theory finds support, while existing theories fail to explain the results. I conclude with a discussion of how moral exemplification theory improves our understanding of party stereotypes and the effects of politicians’ position-taking.

**Existing Theories of Trait Perceptions**

Character is a natural way for citizens to evaluate politicians, as trait judgments are made constantly in everyday life (e.g., Rahn et al. 1990). As a result, trait judgments are ubiquitous, requiring little political
sophistication from citizens (e.g., Pierce 1993). These features have led scholars to describe trait impressions as the “central component” in candidate evaluations (McGraw 2003, 389), working as the path through which political information affects candidate evaluations (e.g., Rahn et al. 1990; Druckman and Parkin 2005). Numerous studies demonstrate the impact of political information on trait perceptions, whether from presidential debates, speeches, or news coverage (e.g., Benoit, Hansen, and Verser 2003; Fridkin et al. 2007; Fridkin and Kenney 2011). Overall, trait impressions are crucial to understanding how political information affects candidate evaluations, but we know less about how information (such as candidates’ issue stances) affects trait impressions.

Existing research provides evidence that politicians’ issue stances influence perceptions of their character traits, but the field lacks a consistent explanation for why this should be the case. According to one perspective, issue stances affect trait perceptions largely through a process of cognitive dissonance (e.g., Peterson 2005). Individuals rate politicians’ character traits more negatively when they disagree with their issue stances and more positively when they agree with their stances. This creates a straightforward prediction about the relationship between issue stances and traits, but assumes all traits are equally relevant to any given issue stance. Issue stances do little more than contribute to a global evaluation of politicians.

An alternative approach is based on the psychology of person perception, and has become associated with the trait ownership literature (Hayes 2005; Goren 2007). According to this perspective, people naturally make trait inferences from others’ behavior, and do so effortlessly and automatically (Rahn et al. 1990; for a review, see Uleman et al. 2007). Importantly, people interpret behaviors in terms of the specific trait concepts that they exemplify (Srull and Wyer 1989, 60). For example, symbols (e.g., a rose) linked to individuals cause specific trait impressions (e.g., romantic), but do not generalize to other

---

2Notably, multiple studies experimentally confirm the causal effect of trait impressions on candidate evaluations (e.g., Huddy and Terkildsen 1993; Funk 1996).
traits (Carlston and Mae 2006). According to the trait ownership approach, upon learning a politician’s issue stance, individuals make inferences on the basis of traits that are most widely associated with the issue (Rapoport, Metcalf, and Hartman 1989). For example, even if a person disagrees with a politician’s support for welfare programs, he or she will rate the politician as more compassionate as a result of the association between the welfare and compassion.

The cognitive dissonance and ownership theories make very different predictions about the issue-trait relationship. According to the cognitive dissonance approach, issue stances have equivalent effects across trait dimensions, which obviates the need to distinguish between trait dimensions. According to the ownership approach, citizens make the same trait inferences from a politician’s issue stance, regardless of whether they agree with that stance. Neither account seems complete and the empirical evidence is inconsistent. Researchers often average across issue stances to create an index of issue agreement, and average across trait dimensions to create a trait index, obscuring any heterogeneity in the effects of issues stances on trait impressions (e.g., Rahn et al. 1990; Peterson 2005; Druckman and Parkin 2005). Research that has directly examined issue-trait linkages finds individuals make different inferences from the same information, but provides little insight as to why (Rapoport, Metcalf, and Hartman 1989). As a result, we have little evidence for the assumptions underlying trait ownership theory or for understanding candidates’ strategic use of issue stances.

A Theory of Moral Exemplification

A growing body of work shows that morality forms the basis for judgments about the behaviors of oneself and others (Rai and Fiske 2011) and is dominant in impression formation (Fiske, Cuddy, and Glick 2007). When asked to evaluate someone, individuals seek out information about moral traits (e.g., empathetic, loyal) before competence (e.g., intelligent, hard-working) or sociability (e.g., warm, friendly; Brambilla et al. 2011), and weigh moral traits more heavily in global evaluations and presidential

\[^3\]The authors characterize the results as “idiosyncratic” and call for further research into the question.
approval (Bazinska and Wojciszke 1996; Wojciszke, Bazinska, and Jaworski 1998). On this view, it is more important to know the intentions of others before knowing their ability to carry out those intentions (Fiske, Cuddy, and Glick 2007). Taken together, existing research implies that theories of morality are crucial to understanding how people evaluate candidates.

The defining features of morality make it uniquely relevant to understanding trait impressions. Although the moral foundations (described further below) have some conceptual overlap with personal and political values (e.g., Feldman 1988, Goren 2005), not all values are moral values (Graham et al. 2011), which arguably represent the “deepest” level of values. Most relevant to theories of trait impressions, moral judgments are characterized by their universalism—a tendency to universally apply judgments of right and wrong, regardless of the actor’s own beliefs, culture, or religion (e.g., Haidt, Koller, and Dias 1993; Skitka, Bauman, and Sargis 2005). In other words, while we may acknowledge that others have different values and goals than ourselves, we perceive moral claims as statements of fact. Perceived moral violations thus generate strong emotional reactions and “lower thresholds for harsh dispositional attributions” towards norm violators (Tetlock et al. 2000, 855). Overall, the features of morality that distinguish it from broader values make it ideal for explaining character trait inferences (see also Pizarro and Tannenbaum 2012).

**Broadening the Moral Domain**

Political scientists have spilt much ink on morality, but typically focus on a narrow slice of morality, providing little in the way of a broad theoretical framework. Much of the literature has focused on analyzing issues such as gay rights, stem cell research and abortion as a distinct class of “morality policies” (e.g., Mooney 2001; Grummel 2008). In this sense, morality is often conflated with specific moral beliefs typically held by religious conservatives and many scholars implicitly or explicitly define morality in terms of sin or religion (e.g., Haider-Markel and Meier 1996; Mooney and Lee 1999).
However, moral judgments extend well beyond this narrow class of policies. According to Moral Foundations Theory, moral beliefs can be categorized into five broad domains: Care, Fairness, Authority, Loyalty and Sanctity (Haidt and Joseph 2004; Graham et al. 2011). Each foundation represents a psychological system, with its own evolutionary history, that gives rise to moral intuitions. These intuitions, or automatic flashes of approval or disapproval, form the foundations for cultures to build moral virtues (Haidt and Graham 2007). Under this view, a virtue represents a tendency to show the morally praiseworthy response to a particular type of social situation. For example, compassion can be understood as showing the proper emotional and behavioral response to the suffering of others (Haidt and Joseph 2004). In general, the five moral foundations describe the types of information that we find morally relevant, and the character traits or virtues we use to evaluate others’ behavior.

The first two foundations (Care and Fairness) are the most familiar to American culture, tend to be more strongly endorsed by liberals, and emphasize the individual. The Care foundation reflects our sensitivity to the suffering of others, and associated moral virtues include kindness and compassion. The Fairness foundation is related to the evolutionary idea of reciprocal altruism, and generates ideas of rights and justice. Associated virtues include honesty and impartiality.

The remaining three foundations (Authority, Loyalty, and Sanctity) emphasize community over the individual and are more strongly endorsed by conservatives. The Authority foundation emphasizes support for a hierarchical social structure. Relevant virtues include respect and leadership. Loyalty represents individuals’ propensity to prioritize in-group members over out-group members. Virtues include loyalty and self-sacrifice. Lastly, Sanctity has origins in the emotion of disgust, and concerns keeping the body and spirit free from contamination. Relevant virtues include cleanliness and abstention.

Each moral foundation can be understood as a dispositional sensitivity to particular features of the social environment, such as suffering or disorder, and the corresponding moral obligations. Cultures
build virtues upon these foundations, which reflect an individual’s propensity to uphold a particular foundation. As a result, environmental features that trigger an intuition, or gut response, particularly among those high in the relevant foundation, will also trigger associated virtues or character traits. Given the greater frequency of activation of the foundation and relevant trait concepts among individuals high in a moral foundation, these trait concepts will be more cognitively accessible.

**Accessibility Hypothesis**: trait concepts representing a specific moral foundation are more accessible to those high in the foundation than those low in the foundation.

For example, character traits such as kind and compassionate should be more accessible among individuals high in the Care foundation than those low in the foundation. In turn, greater accessibility of a trait should increase the likelihood of using it to evaluate the behavior of others (e.g., Fazio and Williams 1986; Narvaez et al. 2006).

Not only do moral foundations shape the types of traits we use to evaluate others, but they also influence our broader political attitudes. Indeed, moral foundations strongly predict political ideology (Graham, Haidt, and Nosek 2009; Weber and Federico 2012) and attitudes towards social groups (e.g., environmentalists; Graham et al. 2011) that are said to underlie ideology (e.g., Brady and Henry 1985; Conover and Feldman 1981). Most importantly, moral foundations are strong predictors of a variety of issue stances, even after controlling for political ideology and demographic variables (Koleva et al. 2012). For example, the Care foundation is a strong predictor of opposition to the death penalty, while the Authority foundation is a strong predictor of support for the policy.

Given that political attitudes are influenced by our moral foundations, politicians’ issue stances provide an opportunity to assess which moral foundations they uphold, and thus their character traits. Yet, just as features of particular issues may invoke different moral concerns for different individuals (e.g., Care vs. Authority), issue stances may also elicit different character trait judgments. For example, if an aspect of an issue triggers Care concerns among individuals for whom these concerns are accessible,
it should also trigger the related trait concepts of compassion and sympathy. At the same time, a feature of that same issue might trigger concerns about Authority (e.g., maintaining order) among those for whom Authority concerns are accessible, activating related trait concepts of leadership and respect. Thus, the moral concerns that cause individuals to take different stances on an issue will also cause them to associate different character traits with the issue. Given that more accessible trait concepts are more likely to be used to interpret behavior (Srull and Wyer 1989), individuals will make trait inferences associated with their own moral motivations on the issue. This view fits with a recent argument that our inability to understand moral motivations conflicting with our own leads to a “moral empathy gap” in which moral disagreement is attributed to others’ moral deficiencies (Ditto and Koleva 2011). This leads to the following two hypotheses:

**Agreement Hypothesis:** issue agreement will cause more favorable ratings of a politician on character traits that correspond with the subject’s moral motivations on the issue.

**Disagreement Hypothesis:** issue disagreement will cause less favorable ratings of a politician on character traits that correspond with the subject’s moral motivations on the issue.

These two hypotheses distinguish moral exemplification theory from the ownership approach, which predicts that everyone will make the same trait attributions (e.g., compassionate), regardless of whether they favor or oppose the policy.

The previous two hypotheses are specifically about traits representing a subject’s moral motivations, but the hypotheses are agnostic about the effects of issue stances on traits that are not representative of a subject’s moral motivations. One might expect that an increase in the morally

---

4 Note that while the Agreement and Disagreement hypotheses are mirror images of one another, it is possible that one is supported while the other is not. For example, if moral failures are more diagnostic of character than moral success, moral disagreement may affect trait impressions while moral agreement does not.
relevant trait will cause a halo effect, in which individuals perceived positively (negatively) on one trait dimension are more likely to be perceived positively (negatively) on other trait dimensions (e.g., Funk 1997). However, if effects on other traits are merely halo effects, the effect should be strongest on relevant moral traits.

**Motivations Hypothesis:** the effect of issue agreement/disagreement will be greater for traits that represent the subject’s moral motivations than for traits that do not.

This final hypothesis distinguishes moral exemplification theory from the cognitive dissonance approach to issue-trait connections. If the cognitive dissonance approach is correct, then a politician’s issue stance should exert equal effects across trait dimensions. However, moral exemplification theory predicts that the effect of a politician’s issue stance will be strongest for the morally relevant trait dimension.

In summary, a growing literature provides a rich foundation for understanding how people draw character trait inferences, and how politicians might shape these perceptions through strategic position-taking. According to moral exemplification theory, there should be considerable heterogeneity both in the types of traits individuals use to evaluate politicians, and the trait inferences that individuals make from any given issue stance.

**Study 1**

As a first step in testing the theory outlined above, it is important to establish that moral foundations correspond with the types of traits used to evaluate politicians. In order to test this first hypothesis, an online survey was fielded in September 2011 that recruited subjects through Amazon’s Mechanical Turk (MTurk), an approach that is becoming increasingly common in political science (e.g., Arceneaux 2012). While MTurk does not provide a nationally representative sample, research shows that it is much more diverse than typical convenience samples, and provides high-quality data for low cost (Berinsky, Huber, and Lenz 2012). Subjects were recruited for a survey on “public opinion” and were paid $0.40 for participation, with 274 subjects completing the study.
At the beginning of the survey, subjects were asked to list ten character traits that describe their most admired politician, then ten traits that describe their least admired politician. The only question preceding the open-ended questions asked for the subject’s age, thus there are no concerns that subjects were primed to think about moral beliefs, or certain types of traits. Trait-listing tasks such as this one have been used to measure the chronic accessibility of a trait construct (Higgins, King, and Mavin 1982; Narvaez et al. 2006). Connecting this research to the Accessibility Hypothesis, moral traits should be more chronically accessible for individuals who score high on the relevant moral foundation, and thus should be more prevalent in open-ended responses.\(^5\)

In order to code the open-ended responses as relevant to a specific moral foundation, a set of coding rules was created based on the Moral Foundations Dictionary developed by Graham, Haidt, and Nosek (2009). The dictionary consists of positive and negative words that correspond to each moral foundation and was designed for broad use (Graham et al. 2012; also see Clifford and Jerit 2013). Each trait response was coded as corresponding to one of the five foundations if the trait was listed in the dictionary, or could be considered a synonym for one of these words. The coding procedure generated six categories: one for each moral foundation (combining both positive and negative traits), and a sixth category representing all other responses.\(^6\) A second person double-coded 236 of the trait words, although this approach is common in psychology, it is possible that respondents’ most and least admired politicians are objectively different on these character traits. In order to ensure that the results below are not driven by this effect, I conducted an auxiliary study holding the politicians constant across respondents. The results suggest that the patterns below are driven by accessibility, rather than objective differences between politicians.

Traits used to describe the most and least admired politicians were combined into the same index because both positive and negative traits should be more accessible for individuals endorsing the relevant moral foundation.
yielding a high level of agreement (92%) and reliability ($\kappa=.87$). Respondents then took part in a brief experiment, which is described in more detail in Study 2. Next, subjects filled out the 30-item Moral Foundations Questionnaire, which assesses respondents’ scores on the five moral foundations (Graham et al. 2011). Finally, subjects answered two questions regarding their ideology and partisanship.

According to the Accessibility Hypothesis, higher individual scores on a moral foundation should increase the likelihood of using traits relevant to that foundation to describe politicians. Overall, nearly one-third of the responses could be classified as relevant to one of the moral foundations, suggesting that moral traits are commonly used to evaluate politicians. Fairness traits were the most common, with subjects listing 2.5, on average (out of 18 responses, on average). The most common Fairness traits were positive and negative variants of trustworthy, dishonest, and fair. Care traits were the next most common, with respondents listing an average of 1.3 traits, such as compassionate, unkind, and caring. Subjects listed 1.1 Authority traits, with common responses including dishonorable, respectful, and strong leader. Subjects listed 0.5 Loyalty traits, on average, such as patriotic, disloyal, and loves America. Finally, Sanctity traits were the least common, with subjects listing 0.4 traits, on average. Common Sanctity traits include integrity, religious, and unfaithful.

In order to analyze the open-ended responses, a count variable was created for each category, indicating the number of relevant traits a subject listed. A negative binomial model was used to predict each count variable using each moral foundation, along with controls for ideology and partisanship. Although all subjects were asked to list twenty traits (ten most-admired, ten least-admired), the survey only required two responses to continue. As a result, only 84% of subjects listed the full twenty traits. 

---

7 It is natural to wonder whether the open-ended trait items influenced response to the MFQ. However, the moral foundations are highly stable over time (Graham et al. 2011), making it unlikely that they would be influenced by the open-ended questions. In contrast, asking the MFQ first would almost certainly affect the accessibility measures.
Thus, the number of each type of trait observed is partially a function of the number of traits listed, or each individual’s ‘exposure’ (King 1998). Accordingly, the exposure feature of the negative binomial model is used to account for this problem, which essentially divides out the number of traits answered.

The results are shown in Table 1, with the coefficients of interest in bold. In the first column, the dependent variable is the number of Care traits listed by the respondent. Of primary interest is the effect of an individual’s score on the Care foundation, which should exert a positive effect on the number of Care traits listed. The coefficient is in the expected direction but falls short of statistical significance ($p=.07$, one-tailed). The second column displays the results for Fairness. As expected, the Fairness coefficient is positive and statistically significant ($p=.02$, one-tailed). The third column displays the results for Loyalty traits. As expected, the coefficient is positive and statistically significant ($p=.02$, one-tailed). The fourth column displays the results for Authority traits. The coefficient on the Authority foundation is small and not in the expected direction. Finally, column 5 displays the results for Sanctity. The coefficient is in the expected direction and statistically significant ($p<.05$, one-tailed).

Although the sample size is small, the overall patterns support the theory. The moral foundations tend to predict relevant traits and in only one case did a moral foundation significantly predict an unrelated trait (Loyalty predicted Authority traits). Overall, the moral foundations are associated with which character traits are most cognitively accessible, and thus most likely to be used in evaluation (Fazio and Williams 1986; Narvaez et al. 2006).

Study 2

The results of Study 1 provide evidence that citizens spontaneously attribute moral traits to politicians, and the type of traits used for evaluation are associated with an individual’s moral foundations. Yet, Study 1 does not provide insight into the process through which those attributions might be formed. To

---

8 I report one-tailed p-values here because I have directional hypotheses.
test hypotheses 2-4, two nearly identical experiments were run focusing on Sam Brownback, the Republican Governor of Kansas. The first experiment was embedded in the survey described above, while the second was a laboratory experiment run in July 2011 utilizing an undergraduate subject pool. Results from the two experiments were highly similar, so I pool the data in order to get more precise estimates of the treatment effects (for a similar approach, see Jerit 2009). Further details and the results from each are displayed in the online appendix.

This study focuses on attitudes towards the death penalty, which, according to previous research (Koleva et al. 2012) and an auxiliary test (shown in the online appendix), are most strongly related to the Care and Authority foundations. Specifically, Care should be the relevant moral motivation for those opposed to the death penalty, while Authority is the moral motivation for supporting the death penalty. Thus, according to the Agreement Hypothesis, those opposing the death penalty (agreeing with Brownback) will rate Brownback more favorably on Care traits. According to the Disagreement Hypothesis, those favoring the death penalty will rate Brownback less favorably on Authority traits. Finally, according to the Motivations Hypothesis, these effects should be strongest for the traits that correspond with an individual’s moral motivations. In order to test this hypothesis, each trait serves as a baseline of comparison for the other trait. Subjects who oppose the death penalty should make stronger (positive) trait attributions for Care traits than Authority traits. Conversely, individuals who support the death penalty should make stronger (negative) trait attributions for Authority traits than Care traits.9

Subjects were all given a short description of Sam Brownback, the Republican Governor of Kansas. While not well known, Brownback was not completely unfamiliar to the public, as he ran for

9 The theoretical argument makes cross-cutting predictions for the Care and Authority foundations, so they offer the sharpest test of the hypotheses. Other trait dimensions were not measured out of concern for respondent fatigue. However, an additional experiment (not reported here) provides evidence that the issue stances only affect predicted trait dimensions.
president in 2008. Although denying any intention to run in 2012, Brownback was still considered to be a potential candidate. In the treatment condition, subjects were informed that Brownback opposes the death penalty, while the control group was given no information about his issue stance. The conditions are shown below, with the treatment in italics.

Sam Brownback is the current Governor of Kansas and has served as a U.S. Representative and Senator for Kansas. Sam Brownback is also a member of the Republican party [and is opposed to the death penalty].

All subjects were informed of Brownback’s partisanship because this information is typically available in the mass media, and excluding it may lead to an overstatement of any treatment effects (McGraw 2011). After receiving the description, all subjects were asked to rate how well four randomly ordered character traits describe Brownback, using the standard ANES measure for trait perceptions. Two traits (kind, compassionate) were selected to represent the Care foundation, while the remaining two were selected to represent the Authority foundation (commands respect, strong leader).

Results

Care traits (kind, compassionate) and Authority traits (strong leader, commands respect) were averaged to create two trait indices ($\alpha = .84$, $\alpha = .77$). The Care and Authority indices were each predicted using a seemingly unrelated regression, which allows a correlated error term across multiple linear models (table shown in the online appendix). This approach allows a test of treatment effects across dependent variables while accounting for the likely correlation between treatment effects. Since the Agreement/Disagreement Hypotheses predict the treatment effect will be moderated by attitudes towards the death penalty, in addition to a dummy variable for treatment condition, death penalty

---

10 However, this study used a five-point (rather than four-point), fully labeled scale.
attitudes and an interaction term are included in the model. The interaction term is significant for both trait indices \((p<.05)\), so the marginal effects of the treatment were estimated for each trait. Marginal effects are shown in Figure 1 among those who strongly favor and those who strongly oppose the death penalty.

The left panel of Figure 1 shows the treatment effects for those who strongly oppose the death penalty (agree with Brownback). Supporting the Agreement Hypothesis, the treatment had a positive effect on Care traits (black circle), which represent supporters’ moral motivation. The size of this effect is .60 points on a 5-point scale, which is substantially larger than the effect of a subject moving from being an independent to a strong partisan (.18). The right panel of Figure 1 shows the marginal treatment effects for those who strongly favor the death penalty (disagree with Brownback). Supporting the Disagreement Hypothesis, the treatment had a negative effect on Authority traits (white circle), which represent death penalty supporters’ moral motivation. Among this group, the treatment decreases ratings on the Authority index by .24 on a 5-point scale, which is nearly as large as the effect of changing a subject’s partisanship from an independent to a strong partisan (.32). These results support moral exemplification theory, but do not support the ownership approach adopted in previous literature, which predicts individuals all make the same trait inferences, regardless of their stance on the issue.

However, the most important test is a comparison of treatment effects across trait dimensions, which is executed with a Wald test of the equality of coefficients. According to the Motivations Hypothesis, underlying this approach is the assumption that individuals share the dominant moral motivation for their issue stance (e.g., all who support the death penalty are motivated by Authority concerns). Any slippage in this assumption serves to decrease the likelihood of observing differential effects across trait dimensions (due to possible heterogeneity in individuals’ moral motivations). As a result, this assumption creates a harder test for the Motivations Hypothesis.
Hypothesis, those who oppose the death penalty will interpret Brownback’s issue stance primarily in terms of Care traits. Consistent with this hypothesis, among those opposing the death penalty the treatment had a larger effect on Care traits than Authority traits (difference = .36, \( p < .01 \)). Turning to those who favor the death penalty, the Motivations Hypothesis makes the opposite prediction, with the treatment having a greater effect on Authority traits than Care traits. Indeed, the treatment effect on the Authority index is significantly larger than the effect on the Care index (difference = .27, \( p < .01 \)). These results again support moral exemplification theory, but cut against the cognitive dissonance approach adopted in previous literature, which predicts equal effects on each trait dimension.

Overall the results show strong support for moral exemplification theory. Among those who support the death penalty, learning that Brownback opposes the death penalty causes these subjects to rate him worse on Authority traits, but not Care traits. Among those who oppose the death penalty, the treatment causes subjects to rate him higher on both Care and Authority traits, but has a significantly larger effect on the Care index. Thus, subjects opposed to the death penalty primarily interpret the issue in terms of Care traits, while those in support of the policy interpret it in terms of Authority traits.

**Study 3**

In order to generalize the findings of Study 2, a second laboratory experiment was run during April 2012. The sample consisted of undergraduates enrolled in political science courses at a large southern university. This study focused on Jim Inhofe and his support of waterboarding and enhanced interrogation procedures. The text of the conditions is show below, with the treatment in italics.

Jim Inhofe is a Republican Senator from Oklahoma who has served in Congress since 1987. Senator Inhofe has been a tireless advocate of government reform, seeking greater accountability and transparency in government spending. [He has long been a vocal supporter of using “enhanced interrogation procedures,” such as waterboarding, on suspected terrorists. Senator Inhofe has been critical of attempts to restrict interrogation practices, and has consistently voted against these restrictions.]
Two key features of this design differentiate it from Study 2. First, the treatment was a party-consistent stance, as Republicans are more likely to favor enhanced interrogation. This sets a higher bar for uncovering a treatment effect, as information that is inconsistent with expectations is more diagnostic than information that is consistent with expectations (Fiske and Taylor 2008). Second, the treatment provided more information on Inhofe’s issue stance, describing him as holding a consistent position over time, and acting on this position. Information about consistency and behavior should increase the treatment effect, as they both increase the diagnosticity of information (Fiske and Taylor 2008).

Results

Previous research shows that the Care foundation strongly predicts opposition to enhanced interrogation and the Loyalty foundation strongly predicts support for enhanced interrogation (Koleva et al. 2012). As a result, the Motivations Hypothesis predicts that opponents of enhanced interrogation will assess Inhofe primarily in terms of Care traits, while supporters of enhanced interrogation will assess Inhofe primarily in terms of Loyalty traits. Subjects rated Inhofe on three Care traits (kind, compassionate, and caring; $\alpha=.84$) and three Loyalty traits (patriotic, loyal, and “one of us”; $\alpha=.70$). Finally, attitudes towards enhanced interrogation are measured using two items ($\alpha=.90$).

Similar to Study 2, each trait index was predicted using a seemingly unrelated regression, with a treatment dummy, attitudes toward enhanced interrogation (EI), and an interaction between the two. As expected, the interaction term is statistically significant in both models ($p<.01$; table shown in the online appendix). In order to unpack the results, Figure 2 shows the marginal effects.

12 For example, a CNN/ORC poll shows that a majority of Republicans (69%) support waterboarding, while a majority of Democrats (56%) oppose it.

13 The two items asked respondents to rate their support for the use of “enhanced interrogation procedures, such as waterboarding, on suspected terrorists,” and their agreement with the statement that “enhanced interrogation procedures, such as waterboarding, are never justified.”
Focusing first on the Disagreement Hypothesis, among those who strongly oppose EI (disagree with Inhofe), the treatment has a large negative effect on Care traits (-1.3, \(p<.001\)). Notably, this effect is substantially larger than the effect of moving the full range of the party identification scale (.23). Although the treatment also had a significant effect on Loyalty traits among opponents (-.76, \(p<.001\)), supporting the Motivations Hypothesis, the effect is significantly larger for Care traits than for Loyalty traits (difference = .56, \(p<.001\)). Turning to proponents of EI (agree with Inhofe), the treatment has a positive effect on the Loyalty index (.65, \(p<.001\)), supporting the Agreement Hypothesis. Again, this effect is larger than moving the full range of the party identification scale (.40). The treatment did not have a statistically significant effect on Care traits (-.07, \(p=.54\)) and, supporting the Motivations Hypothesis, the treatment effect on Loyalty traits was significantly larger than the effect on Care traits (difference = .58, \(p<.001\)).

[Figure 2 about here]

Overall the results bear a striking resemblance to the results in Study 2, supporting all three hypotheses, but conflicting with previous theoretical approaches. Notably, the treatment effects were substantially larger than those in the previous study, in spite of the fact that Inhofe’s stance was consistent with partisan stereotypes. The likely explanation for this result is the more extensive information provided in the treatment, which portrayed Inhofe’s stance as persistent across time and consistent with his behavior. The additional information removes concerns that Inhofe’s stance reflects mere position-taking, rather than his true dispositions. Additionally, in both studies the effects of issue stances were typically larger than the effects of partisanship, cutting against the argument that trait perceptions measure little more than partisan affect (e.g., Bartels 2002).

Discussion

Overall, the results demonstrate consistent support for moral exemplification theory, while previous theoretical approaches fail to explain the results. In addition to strong claims to internal validity, the
preceding studies also satisfy many of the criteria of external validity. Shadish, Cook and Campbell (2002, 83) define external validity as the “extent to which a causal relationship holds over variations in persons, settings, treatments and outcomes.” The three studies varied the population (student vs. diverse online sample), setting (laboratory vs. internet), treatment (death penalty vs. enhanced interrogation) and outcomes (Care, Authority, Loyalty traits), yet all arrived at the same conclusions.

Although the results are strong, two issues deserve discussion. First, while the experiments demonstrate substantial treatment effects, in the real world minimal citizen awareness of politics may prevent these effects from being actualized. However, extensive research demonstrates that the issue focus of campaigns and the media shape citizen knowledge of the dominant issues (Kahn and Kenney 2001), and greater coverage of issues in the media leads to a greater impact of issues on vote choice (Kahn and Kenney 1999). Moreover, exposure to information outside of conscious awareness can affect trait perceptions (Bargh and Pietromonaco 1982), suggesting memory of an issue stance is not necessary for the stance to affect trait perceptions. Even mere mentions of Senators’ character traits in the media increases both citizens’ willingness to rate Senators’ traits and the valence of those ratings (Fridkin and Kenney 2011). Thus, there is good reason to believe that politicians are able to influence public trait perceptions through the issue stances they take.

Second, moral exemplification theory does not make predictions about competence traits. However, there is reason to think that perceptions of competence will be influenced by perceptions of moral traits. Indeed, research shows that the perceived efficacy, or costs and benefits of a policy, are affected by moral agreement with the policy (Liu and Ditto 2012). Thus, if a policy stance implies good moral character, it may also imply competence. Reflecting the fact that moral traits tend to dominate person perception (Bazinska and Wojciszke 1996; Wojciszke, Bazinska, and Jaworski 1998; Brambilla et al. 2011), perceptions of competence may be influenced by perceptions of moral character.

Conclusion
Character traits represent a person’s moral dispositions. Individuals showing a strong reaction to the suffering of others are considered compassionate, while individuals sacrificing themselves for the team are considered loyal. Upon witnessing the speech or behavior of a politician, citizens have the opportunity to ask “does this politician share my moral beliefs?” Yet, just as there is heterogeneity in the moral beliefs that shape individuals’ attitudes on a given issue, there also is heterogeneity in the trait information citizens draw from a politician’s issue stance. Across two experiments, the results consistently show that individuals make trait attributions that reflect their own moral motivations for their stance on the issue. These studies are the first to demonstrate that individuals make different types of trait attributions from the same information about an issue stance. As a result, moral exemplification theory holds important implications for our understanding of partisan stereotypes and campaign strategies.

According to trait ownership theory, trait perceptions are engrained in partisan stereotypes (Hayes 2005). This theory is based on an intuitive argument regarding the connection between issues and traits (e.g., social welfare and compassion, foreign policy and leadership), and draws on the existing evidence regarding these connections. Moral exemplification theory provides a deeper explanation for these findings, holding that partisan divides on issue stances and trait impressions are both rooted in moral differences. Liberals rely primarily on the Care and Fairness foundations (Graham, Haidt, and Nosek 2009), explaining Democrats’ trait ownership of Care traits (e.g., compassion). Conservatives, on the other hand, also rely on Authority, Loyalty and Sanctity foundations (Graham, Haidt, and Nosek 2009), explaining Republicans’ ownership of Authority traits (e.g., leadership) and Sanctity traits (e.g., integrity). This study also adds to existing theory, positing that Democrats should own Fairness traits (e.g., tolerance) and Republicans should own Loyalty traits (e.g., patriotism), claims that are yet to be empirically explored. Finally, my argument amends trait ownership theory, suggesting trait ownership is not created merely by the issues parties emphasize, but also the stances parties take on those issues.
Recognizing that individuals on different sides of the political spectrum might not only make trait attributions of a different valence (i.e., positive or negative), but also of a different dimension (e.g., leadership or empathy), adds to our understanding of polarization. While the debate over the extent of mass polarization is ongoing (Hetherington 2009), there is a clear pattern of polarization in the trait evaluations of the parties and presidential candidates (Hetherington and Long 2012; Iyengar, Sood, and Lelkes 2012). One explanation for this finding is increased sorting over time, in which the political parties increasingly consist of individuals sharing the same values and issue stances (Bafumi and Shapiro 2009). Increased partisan sorting suggests Democrats, who rely more on the Care and Fairness foundations, are more likely to perceive Republicans as uncompassionate and prejudiced. Meanwhile, Republicans, who rely more on the Loyalty, Authority and Sanctity foundations, are more likely to perceive Democrats as unpatriotic, subversive and sinful. Thus, moral exemplification theory, in combination with increased partisan sorting, offers a mechanism for the increasing polarization in trait perceptions even in the absence of partisan bias in candidate evaluations.

Turning to elite behavior, campaigns have long engaged in the “dovetailing” of issue stances and character traits in advertisements under the assumption that issue stances help shape trait impressions (e.g., Kern 1989; Just et al. 1996; Kahn and Kenney 1999). Conventional wisdom suggests that stances on issues like the death penalty have powerful effects on trait perceptions. The present study not only solidifies this claim, but also shows that certain issue stances are more effective at influencing some character traits rather than others. Supporting the death penalty improves leadership credentials, but undermines compassion; supporting the torture of suspected terrorists creates an impression of patriotism, but also undermines compassion. As a result, politicians seeking to bolster their image through issue stances must weigh the potential gains and losses on different trait dimensions among various segments of the public. From this one might conclude that the best strategy is to take only popular positions. However, from the standpoint of displaying moral character, taking a widely endorsed
issue stance is unlikely to be successful. Attribution research demonstrates that the more common and expected a particular behavior, the less diagnostic it is about an actor’s dispositions (e.g., Ybarra 2002). To the extent that trait impressions mediate the effects of issue stances (e.g., Rahn et al. 1990; Druckman and Parkin 2005), taking overwhelmingly popular issue stances may do little to bolster a candidate’s favorability.

Additionally, the effect of a particular trait dimension on favorability is conditional on political context. For example, the threat of terrorism increases the emphasis on leadership traits (Merolla and Zechmeister 2009; Berinsky 2009). Likewise, poor economic conditions have been shown to benefit Democratic politicians (Merolla and Zechmeister 2013), perhaps due to greater weighting of compassion traits. That said, politicians are not held hostage by political context. Substantial evidence demonstrates that elites can prime particular trait dimensions through speeches and news coverage (e.g., Druckman 2004; Druckman and Holmes 2004; Druckman, Jacobs, and Ostermeier 2004). Going forward, moral exemplification theory may illuminate how issue salience influences the weighting of specific trait dimensions and contributes to partisan electoral advantages.
References


Biographical Statement:

Scott Clifford is a Postdoctoral Fellow at Duke University, Durham, NC 27707.
Table 1. Open-Ended Trait Responses as a Function of Subjects' Moral Foundations

<table>
<thead>
<tr>
<th>Character Traits:</th>
<th>Care</th>
<th>Fairness</th>
<th>Loyalty</th>
<th>Authority</th>
<th>Sanctity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moral Foundations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Care</td>
<td>0.99</td>
<td>-0.43</td>
<td>-0.06</td>
<td>-1.00</td>
<td>-0.83</td>
</tr>
<tr>
<td>Fairness</td>
<td>0.69</td>
<td>0.88*</td>
<td>1.17</td>
<td>-0.33</td>
<td>1.36</td>
</tr>
<tr>
<td>Loyalty</td>
<td>0.32</td>
<td>0.06</td>
<td>1.95*</td>
<td>1.16*</td>
<td>-1.04</td>
</tr>
<tr>
<td>Authority</td>
<td>0.42</td>
<td>-0.17</td>
<td>-0.79</td>
<td>-0.10</td>
<td>-0.84</td>
</tr>
<tr>
<td>Sanctity</td>
<td>0.22</td>
<td>0.01</td>
<td>0.58</td>
<td>0.09</td>
<td>1.25*</td>
</tr>
<tr>
<td>Controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Party ID</td>
<td>-0.05</td>
<td>0.07*</td>
<td>0.02</td>
<td>-0.03</td>
<td>0.31*</td>
</tr>
<tr>
<td>Ideology</td>
<td>-0.09</td>
<td>-0.04</td>
<td>0.23*</td>
<td>0.06</td>
<td>-0.14</td>
</tr>
<tr>
<td>Constant</td>
<td>-4.01*</td>
<td>-2.32*</td>
<td>-6.63*</td>
<td>-2.63*</td>
<td>-4.61*</td>
</tr>
<tr>
<td>Alpha</td>
<td>0.20*</td>
<td>0.00</td>
<td>0.36*</td>
<td>0.40*</td>
<td>0.12</td>
</tr>
<tr>
<td>Observations</td>
<td>274</td>
<td>274</td>
<td>274</td>
<td>274</td>
<td>274</td>
</tr>
</tbody>
</table>

Results from negative binomial regression with exposure. Alpha represents overdispersion parameter.

* p < .05, one-tailed. Standard errors in parentheses.
Figure 1: Treatment Effects by Death Penalty Attitudes

Note: figure displays the treatment effect on each trait dimension with 95% confidence intervals, as estimated from the model in Table A1. Estimates in the left (right) panel represent effects among those strongly opposing (favoring) the death penalty.
Figure 2: Treatment Effects by Enhanced Interrogation Attitudes

Note: figure displays the treatment effect on each trait dimension with 95% confidence intervals, as estimated from the model in Table A2. Estimates in the left (right) panel represent effects among those strongly opposing (favoring) EI.