The Moral Foundations Hypothesis Does Not Replicate Well in Black Samples

Don E. Davis and Kenneth Rice
Georgia State University

Daryl R. Van Tongeren
Hope College

Joshua N. Hook
University of North Texas

Cirleen DeBlaere
Georgia State University

Everett L. Worthington Jr.
Virginia Commonwealth University

Elise Choe
Georgia State University

The current study examined the generalizability of the moral foundations hypothesis (Graham, Haidt, & Nosek, 2009), which predicts that conservatism will be positively related to the binding foundations (i.e., virtues of ingroup/loyalty, authority/respect, and purity/sanctity). Religiosity has been consistently linked with the binding foundations in predominately White samples, but Black people in the United States are both more religious and more liberal than White people. In a sample of college students ($N = 693$; 58.3% Black, 41.7% White), examination of measurement invariance suggested metric, but not scalar invariance. The relationship between conservatism and the binding foundations—specifically, respect/authority and purity/sanctity—was weaker in Black people than in White people. These results were replicated in a second sample ($N = 490$; 63.5% Black, 36.5% White) using a 4-item measure of conservatism rather than a single item. Once again examination of measurement invariance suggested metric but not scalar invariance, and conservatism was more weakly related to the binding foundations in Black people than it was in White people. Implications for future theory and research are discussed.

Keywords: moral foundation theory, measurement, invariance, politics

Moral foundations theory (Graham & Haidt, 2010) has quickly become a leading theory in moral psychology, particularly in the study of political identity. Thus far, the centerpiece of the theory has been the moral foundations hypothesis, which predicts that conservatives tend to draw on virtues associated with binding communities (i.e., ingroup/loyalty, authority/respect, and purity/sanctity) more than liberals. Although evidence supporting this hypothesis has come from a variety of samples (Graham et al., 2009; Graham et al., 2011; Koleva, Graham, Iyer, Ditto, & Haidt, 2012), most of these samples consisted of predominately White participants, including the sample used to develop the Moral Foundations Questionnaire (MFQ), which was 87% White (Graham et al., 2011).

We suggest that a conspicuously unanswered question is whether the link between conservatism and the binding virtues holds true in non-White samples and especially Black samples. One reason this link may not hold true concerns the relations between race/ethnicity, political conservatism, and religiosity in the United States. Namely, religiosity, which has been associated with the binding virtues in initial studies (Graham & Haidt, 2010), is generally associated with political conservatism among White people in the United States (Hall, Matz, & Wood, 2010). However, Black individuals tend to be both more religious and more politically liberal than White individuals in the United States (Boyd-Franklin, 2010). For example, Glass and Nath (2006) found that Black women in the United States were substantially more likely to endorse conservative religious beliefs and attend services than were White women (for a review, see Mattis & Grayman-Simpson, 2013), but politically, their views were substantially more liberal. If religiosity is primarily linked with the binding virtues, then Black people ought to score high on binding virtues; but if political conservatism is primarily linked with the binding virtues, then Black people ought to score low on the binding virtues. Something has to give: We wondered whether the moral foundation hypothesis would replicate in Black people.

Overview of Moral Foundations Theory

Moral foundations theory represents a major paradigm shift in moral psychology. Historically, following suit of other academic disciplines (e.g., philosophy or theology), moral psychologists defined morality based on abstract, cognitive principles associated with the protection of individual rights—especially prevention of

harm (Gilligan, 1982) and unjust treatment (Kohlberg, 1969). However, basic research in moral psychology demonstrated that (a) moral judgments are influenced by affective reactions (e.g., disgust and contempt) in addition to abstract cognitive principles; and that (b) many situations other than seeing people harmed or treated unfairly can evoke moral judgments (Haidt, 2001). Haidt's (2001) seminal article on how moral decisions are affectively driven and socially embedded reframed subsequent approaches to understanding how individuals assess the morality of instances, individuals, and institutions.

Moral foundations theory was proposed in response to the need to account for the diversity of moral systems across cultures. The theory sought to distill many virtues into core domains that provide human cultures with a first draft of a moral system upon which societies elaborate and contextualize. The original theory suggested that there are five moral foundations that align with systems that evolved to address key social problems (e.g., regulation of attachment, altruistic reciprocity).

These five domains fall into two broader categories: individualizing and binding moral domains (Graham & Haidt, 2010). The individualizing moral domains regulate moral judgments associated with protecting individual rights, consistent with the historic focus of moral psychology. The harm/care domain involves basic concerns for preventing suffering of others, and is associated with the attachment system. The fairness/reciprocity domain involves concerns about preventing unfair treatment of others, inequality, and more abstract notions of justice. This domain is associated with the altruistic reciprocity system. In contrast, the binding moral domains include three domains emphasized more strongly in collectivistic cultures and have largely been neglected by prior work in moral psychology. The ingroup/loyalty domain pertains to obligations of group membership, such as loyalty, self-sacrifice, and vigilance against betrayal. This domain is associated with systems that regulate formation of alliances. The authority/respect domain pertains to protecting the social order and regulating obligations of hierarchical relationships, such as obedience, respect, and proper role fulfillment. This domain is associated with systems evolved to manage social hierarchies. The purity/sanctity domain pertains to social and spiritual contagion, including virtues of chastity, wholesomeness, and control of desires. This domain is associated with systems evolved to prevent disease or social contamination.

To assess these five domains, Graham et al. (2011) developed the MFQ. They evaluated the scale's psychometric properties and evidence of construct validity using a massive sample collected through their website (i.e., yourmorals.org; $N = 34,476$; 37% female; 87% White). Results of exploratory factor analyses indicated a two-factor structure based on parallel analysis, but Graham et al. compared their preferred five-factor model to the two-factor model and decided to retain a five-factor structure based on model comparisons of the two- and five-factor models using confirmatory factor analysis (CFA). Estimates of reliability for the five scores derived from the MFQ have varied widely, with average Cronbach's alpha coefficients ranging from .65 to .84. Graham et al. indicated that relatively lower estimates of internal consistency were less of a concern, because they adopted an alternative approach to scale development that prioritized content breath over reliability. Scores from the MFQ demonstrated initial evidence of construct validity across several studies. For example, the sub-scales were predictably related to values and attitudes toward social groups (Graham et al., 2011), and incrementally predicted a variety of criterion variables above and beyond general values (Schwartz, 1992) and five-factor personality variables (Costa & McCrae, 1992). In addition, the binding moral domains, especially purity/sanctity, incrementally predicted attitudes on 13 controversial political topics (e.g., such as same-sex relationships, abortion, casual sex, flag burning, torture, or stem cell research), above and beyond other covariates (Koleva et al., 2012).

The Moral Foundations Hypothesis

The moral foundations hypothesis (Graham et al., 2009), one of the more provocative predictions of the theory, states that "political liberals construct their moral systems primarily upon two psychological foundations—harm/care and fairness/reciprocity—whereas political conservatives construct moral systems more evenly upon five psychological foundations—the same two as liberals, plus ingroup/loyalty, authority/respect, and purity/sanctity" (p. 1029). Graham et al. (2009) evaluated this hypothesis in four samples using several different methods. In their first study, participants ($n = 1,613$) completed items designed to assess moral relevance (this was an earlier, unpublished version of the MFQ; Graham et al., 2011). As predicted, political liberalism was positively related to the individualizing moral domains and negatively related to the binding moral domains. Furthermore, participants generally rated the individualizing domains higher than the binding domains, and this difference increased at higher levels of liberalism. In their second study, an online sample from the United States ($n = 2,212$) completed explicit and implicit measures of political identity and an explicit measure of the moral domains. The explicit measure included moral relevance items from Study 1, as well as items assessing moral judgments of specific situations, based on an earlier version of the MFQ (Graham et al., 2011). Study 2 findings replicated the general pattern of findings from Study 1. In their third study, Graham et al. recruited another large, international sample ($n = 8,193$). This time participants indicated how much money they would have to get paid to engage in various behaviors associated with violating the five moral domains (e.g., for harm/care, "Shoot and kill an animal that is a member of an endangered species"). Consistent with the Moral Foundation Hypothesis, conservatives required more money than liberals to violate the binding moral foundations. As in the first two studies, participants rated the individualizing domains higher than the binding domains, and this difference increased at higher levels of liberalism. Finally, in their fourth study, they compared sermon transcripts from two denominations known to be politically active and either very politically liberal (i.e., Unitarian) or politically conservative (i.e., Southern Baptists). As predicted, they found that conservative sermons, compared to liberal sermons, were more likely to include language of binding domains and less likely to use language of individualizing domains.

Challenges to Moral Foundations Theory

In an attempt to strengthen the theory, critiques of moral foundations theory have surfaced regarding both the comprehensiveness and the cross-cultural generalizability of the foundations. For example, Iyer, Koleva, Graham, Ditto, and Haidt (2012) recently

examined whether the theory explained the moral profiles of libertarians and concluded that a new moral domain of liberty was needed to explain the moral sensibilities of this group. Similarly, Haidt (2012) alluded to the need to differentiate liberal moral intuitions about justice, focused on preventing social bullies (i.e., inequality and the need to restrain the power of elites), from those of conservatives, focused on preventing social loafers who take advantage of a social system without fairly contributing. Thus, the five moral domains are a strong first foray into operationalizing morality, but it is an ongoing effort to advance a version of the theory that fulfills its promise as a cross-cultural, evolutionary theory of morality.

In addition to the comprehensiveness of the five moral foundations, the generalizability of the moral foundations hypothesis has also been called into question. For example, Janoff-Bulman and Carnes (2013) proposed a model predicting that, generally, conservatives prefer proscriptive virtues (i.e., involving inhibition) whereas liberals prefer prescriptive virtues (i.e., involving activation). They suggested that the MFQ, despite claims of broad content representation, had not adequately sampled prescriptive (i.e., liberal) virtues such as social justice on the binding subscales. Thus, developing a sound measure of moral foundations theory is also an ongoing process.

Empirical research on the generalizability of moral foundations theory is inconclusive. In support of its generalizability, Graham et al. (2011) used a large, international sample to develop the scale. Confirmatory factor analyses indicated that a five-factor model exhibited similar fit across 12 international regions; root mean square error of approximation (RMSEA) values ranged from .04 to .06, and comparative fit index (CFI) values ranged from .73 to .88. Graham et al. (2011) also reported predictable mean differences comparing gender (i.e., women scored higher on harm, fairness, and purity) and Eastern versus Western countries (i.e., Eastern countries had higher means on loyalty and purity). Nevertheless, some findings point to problems in generalizability of moral foundations theory. For example, Davies, Sibley, and Liu (2014) sought to replicate the (a) five-factor structure of the MFQ and (b) moral foundations hypothesis in a New Zealand probability sample ($N = 3,994$). The five-factor model showed similar fit to the findings reported by Graham et al. (2011), with an RMSEA of .06 and a CFI of .82 for the total scale. Also, consistent with the moral foundations hypothesis, conservatism was positively related to the binding virtues. However, conservatism was unrelated to the individualizing foundations, rather than negatively related, as predicted by the moral foundations hypothesis.

One important limitation to prior work is that only a few studies have explicitly tested assumptions regarding measurement invariance, which is a standard and crucial step in cross-cultural research (Chen, 2008). As pointed out by Chen (2008), group comparisons are predicated on the assumption that psychometric properties of a measure are comparable between groups—specifically that the factor structure and loadings are comparable between groups ("weak" invariance) and that the item intercepts are comparable between groups ("strong" invariance). Graham et al. (2009) examined and supported weak invariance across three groups (i.e., the United States, the United Kingdom, and other countries). Similarly, Davies et al. (2014) found results supporting weak invariance by gender in New Zealand. Support for weak invariance indicates that group differences in strengths of associations be-

tween MFQ scores and criterion indicators can be evaluated. Curiously, however, neither study reported tests of "strong" invariance, which would seem to be of primary interest for testing between-groups differences in mean scores on the MFQ. Support for strong invariance indicates that analyses of average group differences can be undertaken and results confidently interpreted without concern that systemic measurement properties might be accounting for any detected group differences. Alternatively, if strong invariance is not supported, group differences could be under- or overestimated, and thus the detected group differences may be a function of measurement artifacts (Chen, 2008). Moreover, although samples in prior studies have been large, they often were online convenience samples of highly educated and overwhelmingly White participants (i.e., 87% White; Graham et al., 2011). Thus, it is critical to provide evidence of measurement invariance before evaluating the cross-cultural generalizability of correlations or mean comparisons implied by the theory.

In a similar vein, it is possible that some of the binding domains may include content that conflates religiosity with the moral construct. For example, the sanctity/purity subscale includes an item that seems to explicitly address religiosity (i.e., "Whether or not someone acted in a way that God would approve of"). Such items may function differently in groups that are both strongly religious and politically liberal. For example, using cluster analysis, Haidt, Graham, and Joseph (2009) identified four clusters. The first two aligned well with the moral foundations hypothesis. Namely, the social conservatives cluster tended to have high scores on all five domains, whereas the secular liberals cluster tended to have high scores on the individualizing but not the binding domains. They also identified two additional clusters. Libertarians had a similar profile to secular liberals, except their means on the individualizing domains were somewhat lower. The religious left had a different pattern, with high scores on the binding foundations and especially high scores on the individualizing foundations. This latter group likely aligns well with many racial/ethnic minorities who are religious. For example, in a recent cluster analysis of 23 political issues, Dimock, Doherty, Kiley, and Vidya (2014) identified four left-leaning groups (i.e., solid liberals, 15%; faith and family left, 12%; next generation left, 12%; and hard-pressed skeptics, 13%) and three right-leaning groups (i.e., 12% steadfast conservatives, 10% business conservatives, and 14% young outsiders). The faith and family left group was the only cluster in which racial/ethnic minority people constituted a majority (i.e., 59%). Thus, it is important to investigate more directly whether the Moral Foundations Hypothesis generalizes to racial/ethnic minorities, because this is a population in which religiosity and political conservatism might not be related in the same way as they are in White people. But before we examine our primary moderation hypothesis, we must first ensure the MFQ supports at least a metric invariance model.

Overview and Hypotheses

We had two primary purposes in the present studies, which were to examine measurement invariance of the MFQ and generalizability of the moral foundations hypothesis in Black people. This was especially important given that racial/ethnic minority individ-

uals were severely undersampled in the MFQ validation sample.[1] We reasoned that investigating a sample of Black people in the United States might provide a particularly potent test of the generalizability of the moral foundations hypothesis and the measurement adequacy of the MFQ. Thus far, support for this hypothesis has come from predominately White samples. Although political conservatism is positively linked with religiosity in White people in the United States, this association does not hold in Black people who tend to be more religious and more liberal than White people. Given the importance of these potentially competing ideologies in shaping social cognition, moral foundation theory may require additional nuance to accurately describe the relationship between conservatism and the binding foundations in Black people. Our primary hypothesis, which we tested in two independent samples, was that conservatism would be more strongly linked with the binding foundations in White people than in Black people. Before testing this hypothesis, we tested our assumption that the MFQ had at least weak invariance by race/ethnicity, although we also examined assumptions of strong invariance.

Study 1

Method

Participants and procedure. Participants were 693 students (472 women; 220 men) recruited from undergraduate courses at a large public urban university in the Southeastern United States. The mean age was 24.24 (SD = 6.36). The sample was 58.3% Black and 41.7% White and was predominately religious/spiritual (87.7%). After providing consent online, participants completed measures of political conservatism and moral foundations.

Measures. Conservatism was assessed with a single item, which asked "How conservative are you politically?" on a scale ranging from 0 (*very liberal*) to 100 (*very conservative*). The five moral foundations were assessed with the 30-item MFQ (Graham et al., 2011). Participants rated the moral relevance of 15 statements (e.g., Whether or not someone suffered emotionally) on a 6-point scale, ranging from 0 (*not at all relevant*) to 5 (*extremely relevant*). In addition, they rated their degree of agreement with 15 items related to moral judgments (e.g., Men and women have different roles to play in society) on a 6-point scale ranging from 0 (*not at all relevant or strongly disagree*) to 5 (*extremely relevant or strongly agree*). The MFQ has five subscales, and high scores on each of the five subscales indicate a high priority for that particular moral domain. As described earlier, the subscales showed evidence of reliability, with average Cronbach coefficients alpha ranging from .65 to .84. In addition, the subscales showed evidence of construct validity, being related to political identity and stances on controversial topics (Graham et al., 2009). Furthermore, the subscales predicted criteria above and beyond a measure of values (Graham et al., 2011). In the White sample, the coefficients alpha for the current sample were .64 for harm/care, .61 for fairness/reciprocity, .63 for ingroup/loyalty, .58 for authority/respect, and .65 purity/sanctity. In the Black sample, the coefficient alpha were .69 for harm/care, .68 for fairness/justice, .66 for ingroup/loyalty, .66 for authority/respect, and .76 for purity/sanctity. Although these coefficient alphas are somewhat low, they are consistent with prior studies.

Results and Discussion

We ran a series of measurement and substantive multigroup models with Mplus 7.11 (Muthén & Muthén, 1998–2012) using ML estimation. The two groups compared were Black people and White people. Items of the MFQ were used as indicators of the harm/care, fairness/reciprocity, ingroup/loyalty, authority/respect, and purity/sanctity factors, which were modeled as correlated factors. Several fit indices were examined to evaluate the overall fit of the model—the chi-square value, CFI, the square-root-mean-residual (SRMR), and RMSEA. As a rule of thumb, a CFI around .95, an SRMR equal or less than .08, and an RMSEA equal or less than .06 suggest good fit (Hu & Bentler, 1999). We also conducted chi-square difference tests of nested models and examined change in CFI (Kenny, 2014).

Before testing our primary hypothesis, we conducted confirmatory factor analyses to examine invariance assumptions. Accordingly, we tested a series of models including the configural model that freely estimated factor loadings in both groups, the metric or "weak" invariance model that constrained the factor loadings to be the same in both groups, and the scalar or "strong" invariance model that constrained the intercepts to be the same in both groups. We found that the configural model with all five factors estimated simultaneously showed poor fit, χ^2(790) = 3476.72, p > .001, CFI = .59, RMSEA = .10, SRMR = .11. Because multiple factor models may overly complicate invariance testing (Bontempo & Hofer, 2007), we examined each factor and corresponding item subset separately (see Table 1). With the exception of the ingroup/loyalty subscale, the subscales reasonably supported the assumption of metric invariance implied by our subsequent models. However, none of the scalar invariance models could be supported given substantial and significant decrements in fit when constraining intercept invariance between the groups (Mplus output with global fit and estimates of intercepts are available upon request from first author).

Our primary hypothesis was that the association between political conservatism and the binding moral foundations would be weaker in Black people than in White people. To test this hypothesis, we examined a series of multigroup models in which we regressed each moral foundations subscale on conservatism scores. We estimated the model in Black people and White people separately, and then we examined a model in which the path between conservatism and the moral foundation was constrained to be equal across groups. Results are presented in Table 2. Chi-square difference tests indicated no significant difference between the freely estimated and constrained models for harm/care or fairness/reciprocity; a significant difference was obtained for authority/respect and purity/sanctity. Thus, the relationship between political conservatism and harm/care, $\Delta\chi^2$(1) = .03, p = .862, and with fairness/reciprocity was similar across groups, $\Delta\chi^2$(1) = .07, p = .791. We did not examine ingroup/loyalty due to lack of weak invariance. The relationship between political conservatism and authority/respect was weaker in Black people (β = .25) than it was in White people, β = .47, $\Delta\chi^2$(1) = 3.92, p = .048. The relation-

---

[1] We requested the dataset from Jesse Graham who graciously provided the data immediately. Of around 111,568 who reported race/ethnicity, only around 1,662 identified as Black/African-American; of these, only around 277 reported attending religious services at least once per month.

Table 1
Invariance Testing in Study 1

Model	df	χ^2	$\Delta\chi^2$	p	CFI	RMSEA	95% CI	SRMR	Interpretation
Harm configural	18	111.81	—	—	**.71**	**.12**	[.10, .15]	.07	Questionable
Harm metric	23	120.49	8.68	.123	**.70**	**.11**	[.09, .13]	.08	Pass
Harm scalar	29	258.49	138.00	**.000**	**.30**	**.15**	[.14, .17]	**.13**	Fail
Fairness configural	18	53.16	—	—	.95	.08	[.05, .10]	.04	Adequate
Fairness metric	23	71.21	18.05	**.003**	.93	.08	[.06, .10]	.07	Fail
Fairness scalar	29	144.19	72.98	**.000**	**.83**	**.11**	[.09, .13]	**.09**	Fail
Loyalty configural	18	93.97	—	—	**.91**	**.11**	[.09, .13]	.05	Adequate
Loyalty metric	23	111.3	17.33	**.004**	.89	**.11**	[.09, .13]	.08	Fail
Loyalty scalar	29	163.27	51.97	**.000**	**.84**	**.12**	[.10, .13]	**.10**	Fail
Authority configural	18	196.72	—	—	**.68**	**.17**	[.15, .19]	.08	Questionable
Authority metric	23	203.33	6.61	.251	**.68**	**.15**	[.13, .17]	.09	Pass
Authority scalar	29	302.38	99.05	**.000**	**.51**	**.17**	[.15, .18]	**.12**	Fail
Purity configural	18	71.23	—	—	.89	**.09**	[.07, .12]	.06	Adequate
Purity metric	23	76.8	5.57	.350	.89	.08	[.06, .10]	.06	Pass
Purity scalar	29	119.54	42.74	**.000**	**.81**	**.10**	[.08, .11]	.07	Fail

Note. Bold type is an indication of potentially poor fit. CFI = comparative fit index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standard root mean residual.

ship between political conservatism and purity/sanctity was weaker in Black people ($\beta = .35$) than it was in White people, $\beta = .65$, $\Delta\chi^2(1) = 4.45$, $p = .035$.

Taken together, results from Study 1 suggest that the Moral Foundations Hypothesis had stronger support for White people than it did for Black people. In addition, CFA results suggest concern, especially regarding the ingroup/loyalty factor. This subscale had especially poor fit even when factor loadings were estimated freely in the configural model. In addition, across subscales, there were problems with scalar invariance, which suggests that researchers may need to carefully consider whether this scale is working similarly across groups before conducting mean comparisons. A limitation of our study, however, was that we assessed conservatism with a single item, and we did not directly assess religious commitment. Assessing religious commitment would allow us to confirm that relationship between conservatism was different for Black people than for White people. The second study was designed to address those limitations.

Study 2

The purpose of the current study was to replicate the findings of Study 1. We also included two improvements: a four-item measure of conservatism was used instead of a single item and a measure of

religious commitment (Worthington et al., 2003) was added to test whether the relationship between religious commitment and conservatism was different between Black people and White people in our sample, which was assumed in our justification for examining the moral foundations hypothesis in Black people. Otherwise, the measurement and substantive hypotheses were identical to those posed in Study 1.

Method

Participants and procedure. Participants were 490 students (345 women; 145 men) recruited from undergraduate courses at a large public university in the Southeastern United States. The mean age was 24.79 ($SD = 7.13$). The sample was predominately religious/spiritual (86.7%). The sample was 63.5% Black and 36.5% White.

Measures. As in Study 1, we assessed moral foundations with the MFQ (Graham et al., 2009). In the White subsample, coefficients alpha for the current study were .59 for harm/care, .61 for fairness/reciprocity, .72 for ingroup/loyalty, .67 for authority/respect, and .79 for purity/sanctity. In the Black subsample, the Cronbach's coefficient alpha were .69 for harm/care, .63 for justice/reciprocity, .68 for ingroup/loyalty, .65 for authority/respect, and .73 for purity/sanctity. Political conservatism was assessed

Table 2
Examination of Whether Race/Ethnicity Moderates the Relationship Between Political Conservatism and the Moral Foundations in Studies 1 and 2

	Study 1						Study 2						
	Blacks			Whites			Blacks			Whites			
	b	SE	p	b	SE	p	b	SE	p	b	SE	p	Interpretation
Harm/care	−.06	.07	.333	−.05	.07	.457	−.11	.07	.108	−.04	.09	.707	No difference
Fairness/justice	−.17	.07	.010	−.11	.07	.114	−.16*	.07	.027	−.21	.09	.022	No difference
Ingroup/loyalty	—			—			—			—			Not tested due to poor fit
Respect/authority	.25	.07	.000	.46	.06	.000	.00	.07	.913	.43	.08	.000	Weaker in Blacks in both samples
Purity/sanctity	.35	.06	.000	.61	.05	.000	.12	.07	.074	.60	.06	.000	Weaker in Blacks in both samples

with four items (alpha for the Black sample was .80 and for the White sample was .87), including the item used in Study 1, and the following three items (i.e., foreign policy issues, economic issues, and social issues), which participants rated on a 5-point scale ranging from 1 (*very liberal*) to 5 (*very conservative*). We also included the 10-item Religious Commitment Inventory (Worthington et al., 2003) to assess religious commitment. Items are rated on a 5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). An example item is "My religious beliefs lie behind my whole approach to life." Worthington et al. (2003) reported coefficients alpha ranging from .88 to .98. Furthermore, the RCI-10 was significantly and positively correlated with a single-item measure of religiosity, frequency of attendance of religious activities, and self-rated spiritual intensity. The Cronbach's coefficient alpha for the White sample was .95 and was .93 for the Black sample.

Results and Discussion

As in Study 1, we ran a series of multigroup CFAs (i.e., comparing White people and Black people) with ML estimation using Mplus 7.11. Again, items of the MFQ were used as indicators of the harm/care, fairness/reciprocity, ingroup/loyalty, authority/respect, and purity/sanctity factors, which were modeled as correlated. As in Study 1, the configural model with all five factors estimated simultaneously indicated poor fit, $\chi^2(790) = 1869.01$, $p > .001$, CFI = .57, RMSEA = .11, SRMR = .12. Therefore, subsequent analyses were focused on the individual subscales, estimating the configural, metric, and scalar models (see Table 3). Our planned analyses required metric invariance, and this level of measurement support generally held across subscales, with the exception of ingroup/loyalty, which showed very poor fit for both the configural and weak invariance models (see Table 3). Again, we did not find support for scalar invariance.

Next, we examined our primary hypothesis, which was that the relationship between political conservatism and the binding foundations would be weaker in Black people than in White people. As in Study 1, we examined a series of multigroup models that regressed moral foundation subscale factors on the conservatism

factor. After estimating the relationship freely in both samples (i.e., Black people and White people), we examined a model in which the relationship between conservatism and the respective moral foundation was constrained to be equal across groups. Results are reported in Table 2. Chi-square difference tests indicated no significant difference between the freely estimated and constrained models for harm/care and justice/fairness; a significant difference was obtained for authority/respect and purity/sanctity. Thus, race/ethnicity did not moderate the relationship between conservatism and the individualizing foundations of harm/care, $\chi^2(1) = .94$, $p = .317$, or justice/fairness, $\chi^2(1) = .02$, $p < .888$. Consistent with our primary hypothesis, the relationship between conservatism and authority/respect was stronger in White people ($\beta = .43$) than it was in Black people, $\beta = .00$, $\chi^2(1) = 7.04$, $p < .008$. The relationship between conservatism and purity/sanctity was also stronger in White people ($\beta = .60$) than it was in Black people, $\beta = .12$, $\chi^2(1) = 6.59$, $p < .010$. We also examined a multigroup model that regressed religious commitment on conservatism in White people and Black people separately, followed by a model in which this relationship was constrained to be equal across groups. The freely estimated model provided a significantly better fit. As predicted, the relationship between conservatism and religious commitment was stronger in White people ($\beta = .47$, $p < .001$) than it was in Black people ($\beta = .17$, $p = .009$), $\chi^2(1) = 9.19$, $p < .001$.

Taken together, the results of Study 2 replicated those of Study 1. In fact, the effects were stronger in this second study, perhaps due to using a four-item latent construct rather than a single-item to assess political conservatism. Furthermore, the MFQ generally met assumptions of metric invariance, but not scalar invariance.

General Discussion

Moral foundations theory has dramatically shifted ideas within moral psychology over the last decade. In just a few years, the only published measure to operationalize this theory has been used in a large number of studies. Furthermore, although the centerpiece of the theory—the moral foundations hypothesis—has found quali-

Table 3
Invariance Testing in Study 2

Model	df	χ^2	$\Delta\chi^2$	p	CFI	RMSEA	95% CI	SRMR	Interpretation
Harm configural	18	40.85	—	—	.95	.07	[.04, .10]	.05	Adequate
Harm metric	23	46.87	6.02	.304	.95	.07	[.04, .09]	.06	Pass
Harm scalar	29	61.65	14.78	**.022**	.93	.07	[.04, .09]	.07	Fail
Fairness configural	18	54.71	—	—	.93	.09	[.06, .12]	.05	Adequate
Fairness metric	23	60.55	5.84	.322	.92	.08	[.06, .11]	.07	Pass
Fairness scalar	29	124.1	63.55	**.000**	**.81**	**.12**	[.10, .14]	**.10**	Fail
Loyalty configural	18	139.53	—	—	**.77**	**.17**	[.14, .19]	.08	Questionable
Loyalty metric	23	142.86	3.33	.649	**.77**	**.15**	[.12, .17]	**.09**	Pass
Loyalty scalar	29	204.82	61.96	**.000**	**.66**	**.16**	[.14, .18]	**.12**	Fail
Authority configural	18	45.16	—	—	.94	.08	[.05, .11]	.05	Adequate
Authority metric	23	54.05	8.89	.113	.93	.07	[.05, .10]	.06	Pass
Authority scalar	29	69.82	15.77	**.015**	.90	.08	[.05, .09]	.08	Fail
Purity configural	18	119.9	—	—	**.83**	**.15**	[.13, .18]	.07	Questionable
Purity metric	23	129.45	9.55	.089	**.83**	**.14**	[.11, .16]	.08	Pass (marginal)
Purity scalar	29	160.44	30.99	**.000**	**.79**	**.14**	[.12, .16]	**.10**	Fail

Note. Bold type is an indication of poor fit. CFI = comparative fit index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standard root mean residual.

fied support in some samples, the generalizability of this hypothesis has not been explored. Our work focused on whether this hypothesis generalized to Black people in the United States, who are generally more religious and more politically liberal than White people (e.g., Mattis & Grayman-Simpson, 2013). In prior research, the binding foundations (especially the purity/sanctity subscale) have been related to religiosity (Koleva et al., 2012). The samples used to develop the MFQ (Graham et al., 2011) were predominately White, a sample in which conservatism and religiosity tend to be positively related, which is not true in Black people samples. We hypothesized that the binding moral foundations would show a weaker relationship with political conservatism in Black people than in White people.

Across two independent samples, we found support for this hypothesis. For the two binding foundation subscales that met assumptions of weak invariance (i.e., authority/respect and purity/sanctity), the relationship with political conservatism was weaker for Black people than it was for White people. Also in both samples, the ingroup/loyalty subscale showed substantial statistical problems. Two items on this subscale assess loyalty to one's country, one assesses loyalty to one's family, and the others address loyalty more generally. Loyalty toward one's family or country may not be strongly aligned in some subsamples. For example, some individuals may feel very loyal to their family and community but feel lower loyalty toward their country. Furthermore, across subscales, we found consistent violations of scalar invariance, which is assumed whenever a researcher conducts comparison of means between two groups. Although researchers sometimes ignore these assumptions, it is the gold-standard of cross-cultural work to examine measurement invariance before proceeding with theory testing (Chen, 2008). Given that this model aims to articulate a cross-cultural perspective of morality, this is an essential step in evaluating the validity of a measure of moral foundation theory. Indeed, our findings underscore that it is unwise to take these assumptions for granted, because groups in which conservatism and religiosity are weakly related may understand and respond to items differently, which makes sense given that values and virtues are a primary area in which cultures differ.

Limitations and Suggestions for Future Research

Our studies were limited in several ways. First, we used two convenience samples located in the southeastern United States. Thus, our findings provide one clear example in which the moral foundations hypothesis does not replicate well, but future research can explore competing explanations for why the hypothesis was weaker in Black people. Second, our measure of political conservatism, although similar to prior studies on moral foundations theory, is somewhat simplistic. Some scholars conceptualize and measure political orientation as a multidimensional construct, including subdomains such as attitudes toward social and cultural issues (e.g., abortion, gay marriage) and economic issues (e.g., taxation, welfare); thus, it is likely that using a unidimensional measure of political orientation fails to capture the nuanced complexity of other measures of political orientation. We encourage future research to better parse apart the particular dimensions of political orientation that are most strongly aligned with the various moral foundations. A promising strategy might be to use latent class analysis or cluster analysis to examine distinct ideological

groups and then examine mean levels on the various subscales of the MFQ (Graham et al., 2011). Third, the MFQ (Graham et al., 2011), although a starting point for testing the theory, seems to require some revision. Its developers have already alluded to at least two additional dimensions. The findings from our studies suggest that some of the current items may conflate moral foundations with other constructs such as religiosity or racial identity. Most likely, a revised version of the MFQ would benefit from extensive invariance testing related to key identity variables being studied within moral psychology (e.g., gender, race, religion, sexual orientation). Such work will elucidate the ways that identity can affect not just how individuals prioritize virtues, but even how they understand their meaning.

Such Work Is Essential to Realizing the Potential of Moral Foundations Theory

In terms of future theory and research, we think it is critical to bolster the measurement underpinnings of moral foundations theory. Similar to Janoff-Bulman and Carnes (2013), we wonder whether the MFQ sufficiently sampled potential politically liberal expressions of the binding foundations. For example, although liberals in the United States may have a relatively weaker sense of loyalty to country, they may have a stronger sense of loyalty to groups committed to social justice, including for Black people, their racial/ethnic group or their religious/spiritual community. In addition, we wonder whether there are unique domains in which political liberals show a strong deference to tradition. For example, many helping professions (e.g., social workers, clinical/counseling psychologists, counselors) have a strong commitment to multiculturalism and social justice, and within some of these professions, there are still well-established hierarchies and methods of acculturating trainees to the values of the profession. Likewise, there may be some domains in which political liberals have disgust reactions, such as desecration of the environment.

Furthermore, leaving those issues of content validity aside, it may be overstated to say that political liberals draw on binding foundations less than do conservatives. We suggest reframing this hypothesis to emphasize that political liberals and conservatives order the binding virtues differently. Graham et al. (2011) named and scored the binding domains so that higher scores reflected presumably more politically conservative virtues. However, one might just as easily label and score these three constructs so that higher scores reflect more politically liberal virtues, such as diversity (i.e., deemphasizing intragroup loyalty to facilitate strong intergroup alliances), equality (i.e., deemphasizing authority to level social hierarchies), and tolerance (i.e., seeking to deter the use of contempt or disgust to increase ideological conformity). According to this line of thought, instead of implying that liberals only draw on two foundations, it is probably more accurate to say that liberals prioritize different virtues associated with loosening of traditional structures of authority. Accordingly, groups adopt different social strategies (i.e., different calibrations on the five domains) for enhancing cohesion and regulating selfishness for the benefit of the group, and there may be a complex interaction of contextual factors (e.g., need to form alliances with other marginalized groups) that affect the strategy one uses.

Conclusion

Given the value of advancing theories through paradigmatic shifts and measurement development, moral foundations theory has already been a useful framework from which to examine a variety of moral phenomena in various social settings. However, the generalizability of findings across diverse cultures and with diverse participants is an important feature of scientific progress. We hope that our work will spur others to continue to test the boundaries conditions of this theory as a way of refining and strengthening its predictive validity and usefulness in more fully understanding the complexities of human morality.

References

Bontempo, D. E., & Hofer, S. M. (2007). Assessing factorial invariance in cross-sectional and longitudinal studies. In A. D. Ong & M. H. M. van Dulmen (Eds.), *Oxford handbook of methods in positive psychology* (pp. 153–175). New York, NY: Oxford University Press.

Boyd-Franklin, N. (2010). Incorporating spirituality and religion into the treatment of African American clients. *The Counseling Psychologist, 38,* 976–1000. http://dx.doi.org/10.1177/0011000010374881

Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology, 95,* 1005–1018. http://dx.doi.org/10.1037/a0013193

Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment, 4,* 5–13. http://dx.doi.org/10.1037/1040-3590.4.1.5

Davies, C. L., Sibley, C. G., & Liu, J. H. (2014). Confirmatory factor analysis of the Moral Foundations Questionnaire. *Social Psychology, 38,* 1–6.

Dimock, M. D., Doherty, C., Kiley, J., & Vidya, K. (2014). *Beyond red v. blue: The political typology. Fragmented center poses election challenges for both parties*. Washington, DC: Pew Research Center. Retrieved from http://www.people-press.org/files/2014/06/6-26-14-Political-Typology-release1.pdf

Gilligan, C. (1982). *In a different voice*. Cambridge, MA: Harvard University Press.

Glass, J., & Nath, L. E. (2006). Religious conservatism and women's market behavior following marriage and childbirth. *Journal of Marriage and Family, 68,* 611–629. http://dx.doi.org/10.1111/j.1741-3737.2006.00278.x

Graham, J., & Haidt, J. (2010). Beyond beliefs: Religions bind individuals into moral communities. *Personality and Social Psychology Review, 14,* 140–150. http://dx.doi.org/10.1177/1088868309353415

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. http://dx.doi.org/10.1037/a0015141

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology, 101,* 366–385. http://dx.doi.org/10.1037/a0021847

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108,* 814–834. http://dx.doi.org/10.1037/0033-295X.108.4.814

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York, NY: Pantheon/Random House.

Haidt, J., Graham, J., & Joseph, C. (2009). Above and below left–right: Ideological narratives and moral foundations. *Psychological Inquiry, 20,* 110–119. http://dx.doi.org/10.1080/10478400903028573

Hall, D. L., Matz, D. C., & Wood, W. (2010). Why don't we practice what we preach? A meta-analytic review of religious racism. *Personality and Social Psychology Review, 14,* 126–139. http://dx.doi.org/10.1177/1088868309352179

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55. http://dx.doi.org/10.1080/10705519909540118

Iyer, R., Koleva, S., Graham, J., Ditto, P., & Haidt, J. (2012). Understanding libertarian morality: The psychological dispositions of self-identified libertarians. *PLoS ONE, 7,* e42366. http://dx.doi.org/10.1371/journal.pone.0042366

Janoff-Bulman, R., & Carnes, N. C. (2013). Surveying the moral landscape: Moral motives and group-based moralities. *Personality and Social Psychology Review, 17,* 219–236. http://dx.doi.org/10.1177/1088868313480274

Kenny, D. A. (2014). *Measuring model fit*. Retrieved from http://davidakenny.net/cm/fit.htm

Kohlberg, L. (1969). *Stage and sequence: The cognitive-developmental approach to socialization*. New York, NY: Rand McNally.

Koleva, S. P., Graham, J., Iyer, R., Ditto, P. H., & Haidt, J. (2012). Tracing the threads: How five moral concerns (especially purity) help explain culture war attitudes. *Journal of Research in Personality, 46,* 184–194. http://dx.doi.org/10.1016/j.jrp.2012.01.006

Mattis, J. S., & Grayman-Simpson, N. (2013). Faith and the sacred in African American life. In K. I. Pargament, J. J. Exline, & J. W. Jones (Eds.), *APA handbook of psychology, religion, and spirituality: Vol. 2* (pp. 547–564). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/14045-030

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus User's Guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology, 25,* 1–65. http://dx.doi.org/10.1016/S0065-2601(08)60281-6

Worthington, E. L., Jr., Wade, N. G., Hight, T. L., Ripley, J. S., McCullough, M. E., Berry, J. W., . . . O'Connor, L. (2003). The Religious Commitment Inventory–10: Development, refinement, and validation of a brief scale for research and counseling. *Journal of Counseling Psychology, 50,* 84–96.