Moral Foundations and Political Orientation: Systematic Review and Meta-Analysis

J Matias Kivikangas

Department of Information and Service Management, Aalto University, Finland

Quantitative Psychology and Individual Differences, KU Leuven, Belgium

Belén Fernández-Castilla

Faculty of Psychology and Educational Sciences, KU Leuven, Belgium

Imec-ITEC, KU Leuven, Belgium

Simo Järvelä

Department of Psychology, University of Helsinki, Finland

Faculty of Information Technology and Communication Sciences, Tampere University, Finland

Niklas Ravaja

Department of Psychology, University of Helsinki, Finland

Jan-Erik Lönnqvist

Swedish School of Social Science, University of Helsinki, Finland

DOI: 10.1037/bul0000308

Abstract

We investigate the findings that liberals and conservatives rely on different moral foundations. We conducted a comprehensive literature search from major databases and other sources for primary studies that used the Moral Foundations Questionnaire and a typical measure of political orientation, a political self-placement item. We used a predefined process for independent extraction of effect sizes by two authors and ran both study-level and individual-level analyses. With 89 samples, 605 effect sizes, and 33,804 independent participants, in addition to 192,870 participants from the widely used YourMorals.org website, the basic differences about conservatives and liberals are supported. However, heterogeneity is moderate, and the results may be less generalizable than previously thought. The effect sizes obtained from the YourMorals.org data appear inflated compared to independent samples, which is partly related political interest and may be due to self-selection. The association of moral foundations to political orientation varies culturally (between regions and countries) and subculturally (between White and Black respondents an in response to political interest, but not in relation to other demographics). The associations also differ depending on the choice of the social or economic dimension and its labeling, supporting both the bidimensional model of political orientation and the findings that the dimensions are often strongly correlated. Oue findings have implications for interpreting published studies, as well as designing new ones where the political aspect of morality is relevant. The results are primarily limited by the homogeneity of the measures and included studies in terms of sample origins.

Meta-analysis on correlation studies (not experimental)

Keywords: Moral Foundations, political orientation, political ideology, moral psychology, meta-analysis

Public relevance statement:

This study examines the widely published results that liberals and conservatives see morality differently—that they rely on different 'moral foundations'. Our findings suggest that while these differences are mostly stable, they are smaller or more unpredictable outside politically interested White American samples. These differences depend on how the respondents are recruited, from which country and demographic the results are, and how political orientation is measured.

Meta-analysis on the relationship between moral foundations and political orientation

Introduction

Different views on what is and is not moral are a strong factor in contemporary politics. A body of research in the past decade has studied the proposition that differences between the political poles lie not only in disagreements over ideologies, means, or facts but may be grounded in more fundamental differences in the deep psychological feelings of right and wrong. Moral Foundations Theory (MFT) calls the individual characteristics that rule these views and judgments "moral foundations" (Graham et al., 2013; Haidt, 2012). Our meta-analysis focuses on a key finding of research on MFT: that liberals and conservatives differ in the extent to which they endorse each moral foundation.

The titular findings of Graham and others (2009), "Liberals and conservatives rely on different sets of moral foundations", have become increasingly important, as MFT has been widely utilized in the explanation of political differences (e.g., Federico, Weber, Ergun, & Hunt, 2013; Koleva et al., 2012; van der Linden & Panagopoulos, 2019). The rapidly growing body of work that has applied MFT in the context of political research has been based on the analyses and interpretations by Graham and others according to which the original findings are consistent and generalizable across different countries and cultures (Graham et al., 2011). The presumed generalizability influences the interpretations and conclusions of other studies, and researchers rely on these established associations when planning their studies (see e.g., Frimer, Biesanz, Walker, & MacKinlay, 2013; Wetherell, Brandt, & Reyna, 2013; Schein & Gray, 2015). However, an ongoing debate on the reproducibility of psychological science suggests that up to half of the effects that have been reported on and generally accepted may in fact be not replicable (Camerer et al., 2018;

Open Science Collaboration, 2015; Pashler & Wagenmakers, 2012). In the Many Labs 2 study (Klein et al., 2018), which carried out large-scale replications for a number of earlier findings including the association between moral foundations and political orientation, the general patterns were similar but the average effect size was found to be considerably smaller and the effects showed significant heterogeneity, with a portion of the samples showing results opposite to those reported in the original. However, because of the more general nature of the Many Labs 2 project, reasons for these discrepancies were not investigated.

Objectives and focus

The goal of this systematic review and meta-analysis is twofold. First, we examine the extent to which the key findings on moral foundations and political orientation replicate in datasets independent of Graham and colleagues' data. Given the importance of MFT in the field of political research, a more detailed look on the replicability is warranted, as it may be conditional to some currently unknown moderators (Klein et al., 2018; Open Science Collaboration, 2015). Thus, our second goal is to investigate theoretically and empirically grounded moderators that could contribute to the observed heterogeneity in effects, and thereby to provide more reliable and accurate estimates for future research (Gelman & Carlin, 2014; Lakens & Evers, 2014). These include demographics, but also other moderators that are suggested by the empirical studies and methodological critiques.

Our meta-analysis focuses on MFT alone, as no other theory or model of morality has been used to same extent to investigate and explain individual-level political differences (e.g., Federico et al., 2013; Smith et al., 2016; Koleva et al., 2012; Iyer, Koleva, Graham, Ditto, & Haidt, 2012). The current work focuses on political orientation on the common liberal-conservative and left-right dimensions, or social and economic dimensions. Because the details of the analyses were not

planned a priori, we use a transparent exploratory approach in which report all the steps involved in the analyses and publish the data (to the extent we are capable) for independent reanalysis (Wicherts et al., 2016).

MFT and the assumptions about its relationships with political orientation

As opposed to the traditional understanding of morality as rational considerations about justice and/or care (Kohlberg, 1971; Gilligan, 1982; Turiel, 1983; Giammarco, 2016), MFT suggests that conscious moral reasoning is based on intuitive or emotional reactions, which result from multiple different, sometimes conflicting traits (Graham et al., 2013; Haidt, 2012). The established empirical work on these traits has pointed to five moral foundations that are further grouped in two larger categories. Two "individualizing" foundations focus on the welfare of individuals and the responsibility to respect others' rights, comparable to the traditional accounts of morality: *care*[1] (whether someone is hurt or harmed) and *fairness* (whether someone cheats or is deprived of their rights). More controversially, many people also hold that morality concerns issues governed by three "binding" foundations, emphasizing groups and institutions (such as nationalism, traditions, and religion) that keep individuals in check: *loyalty* (whether one's ingroup is betrayed), *authority* (whether respected people, customs, and traditions are properly respected by others), and *sanctity* (whether the intrinsic purity of something is degraded). A critical part of MFT is the idea of moral plurality, that not all foundations are equally important to everyone, but that different people—such as those with different political leanings—endorse these foundations to different extent. More specifically, in the foundational work of Graham et al. (2009), liberals were reported to endorse the individualizing moral foundations care and fairness more than the binding foundations loyalty, authority, and sanctity, whereas conservatives endorsed all five foundations more or less equally,

[1] The names of the moral foundations differ across studies: other name for care has been harm, for loyalty it has been ingroup, and for sanctity, purity. We follow the nomenclature used in Graham et al. (2013).

and each of these at a lower level than at which liberals endorsed care and fairness. This is typically represented in research as negative correlations between conservatism and care and fairness, and positive correlations between conservatism and loyalty, authority, and sanctity.

The above findings were later repeated in the large dataset from the popular YourMorals.org website, which has over the years collected self-reported responses from over 200,000 respondents and contains subsamples of thousands or tens of thousands of respondents from different global regions (Graham et al., 2011). Based on these findings, Graham and colleagues argued that the results are generalizable across cultures and differences in local politics. This claim has had far-reaching consequences for how research in this area is being conducted, and it can be elaborated as the following assumptions, widely accepted in the literature[2]. Assumption 1, Direction: *Conservative political orientation is negatively associated with care and fairness, and positively associated with loyalty, authority, and sanctity.* The first assumption, reflecting the basic findings, guides study design (e.g., Day, Fiske, Downing, & Trail, 2014; Feinberg & Willer, 2015) and informs theory development (e.g., Janoff-Bulman & Carnes, 2013). While there are very few reports of opposite associations, it is possible that some of them may be so close to zero that they require much larger sample sizes to detect than what is normally use in the field, and that they are practically meaningless for many purposes. Assumption 2, Magnitude: *The correlations with political orientation are strongest for authority and sanctity, and weakest for care. (Presumably*

[2] The origin of the assumptions from Graham et al., (2011, p.13, emphasis ours): "The correlations indicate that the liberal-conservative patterns found in the U.S. are robust across national and cultural contexts, both in terms of *direction* (negative correlations [liberals higher] for Harm and Fairness, positive correlations [conservatives higher] for Ingroup, Authority, and Purity) and in terms of *magnitude*: correlations are consistently strongest for Authority and Purity, and weakest for Harm. This suggests that *across cultures*, the most intractable political debates are likely to involve concerns related to respect for traditions/authorities and physical/spiritual purity, while the greatest degree of moral commonality may be found in issues related to harm and care. It also reinforces the claim that political ideology can be self-assessed and that the *unidimensional* left-right construct has some degree of common meaning across societies, despite differences in political party structures and particular national issues […]." We do not intend to beat a strawman: despite being derived from the Graham et al., we do not present these assumptions specifically as theirs. Rather, they are prevalent in the literature, although they are rarely spelled out. The strictest interpretation of each of these assumptions may push them further than may have been intended by Graham et al.

loyalty and fairness lie somewhere in between). This assumption makes the claim that issues about Authority and Sanctity are the most divisive across the political dimension, which may for example guide researchers to focus on these differences instead of those that are assumed to be less divisive.

Assumptions 1 and 2, reflecting the results presented by Graham et al. and commonly viewed to be supported by the extant evidence, are based on less recognized assumptions 3 and 4. These assumptions can be written as follows. Assumption 3, Unidimensionality: *The associations between moral foundations and political orientation not depend on the way in which political orientation is conceptualized and operationalized; different dimensions (such as left-right, and liberal-conservative) have common meaning across contexts.* Assumption 4, Universality: *The associations between moral foundations and political orientation are not dependent on the political culture or other systematic differences between populations*.

The Universality assumption is the most explicit, being one focus of the validation study reported on by Graham and others (2011). They studied the robustness of ideological patterns across cultures and reported some gender differences and differences between East and West in endorsement of the foundations. Although they acknowledged differences between world regions (controlling for some demographics) regarding how moral foundations are related to political orientation, they concluded that the relationships are generalizable to different cultures and populations. Only a few studies have compared the relationships between moral foundations and political orientation in different samples since. For example, Yilmaz and colleagues (2016) reported that the familiar associations replicate in a predominantly Muslim population, and Federico and others (2013) did not find notable North-South differences within the US. Only one study (Davis et al., 2016) has specifically reported on sample differences, and even those differences, between Black and White samples, were still relatively small.

Both the Universality assumption and the Unidimensionality assumption are apparent in the variation in the measurement of political orientation. The original study, with respondents from around the world, explicitly used a one-dimensional approach. Although the scale was anchored only by the labels "liberal-conservative", it included an item description that explained that the dimension was intended to also include what is called "left-right" and other descriptors used in other countries. Subsequently, most researchers have retained the "liberal-conservative" axis but omitted mention of other dimensions. However, some others have used the "left-right" anchors alone, sometimes without any particular justification (Vainio & Mäkiniemi, 2016; van Leeuwen & Park, 2009), sometimes briefly reasoning that it makes more sense in their country (Dawson & Tyson, 2012; Klein et al., 2018; Nilsson & Erlandsson, 2015). This is only reasonable if the authors assume that the relationship between moral foundations and political orientation are the same regardless of the anchors used: that the "left-right" dimension taps into the same underlying political orientation as the "liberal-conservative" dimension, rather than these being two different dimensions. A further variation of this unidimensional approach is to compute a compound political orientation variable even when the measures employed have tapped into several dimensions (e.g., Federico et al., 2013; Wester et al., 2015). Other studies, however, use "left-right" and "liberal-conservative" axes as separate items, indicating a belief that they are not the same, and possibly that each political dimension may have a unique relationship with moral foundations (e.g., Dimdins, Sandgren, & Montgomery, 2016; Kivikangas et al., 2017; Scott & Pound, 2015). Whether various conceptualizations and operationalizations of political orientation are similarly associated with the moral foundations is currently unknown, hampering researchers planning new studies.

Possible differences between different samples and/or political dimensions could help explain some of the heterogeneity found across different studies, especially regarding the effect

sizes. For instance, Graham and others' own studies demonstrate a diverse range of effect sizes: in the 2011 validation study the correlations between sanctity and authority foundations and political orientation averaged .49 and .48 across the international samples (ranging from .42 to .58 for sanctity, and from .37 to .56 for authority), whereas in the original study[3] the corresponding correlations were .26 and .21, respectively—about half the effect size. In independent studies, these correlations are sometimes high (e.g., Black & Reynolds, 2016, *r*s = .52 and .47), sometimes low (e.g., Ritter, 2014, both *r*s = .18). The Many Labs 2 reported an average *r* = .14 for the relationship between conservatism and an aggregate binding MFs (Klein et al., 2018), but their Figure 2 presenting the effect size distributions also shows a number of studies with negative associations. Although these findings are more or less consistent with Assumption 1, the consistency is not very high given the centrality of this assumption. In addition, regarding Assumption 2, Graham and others (2011) claim that sanctity and authority consistently have the strongest relationships to political orientation, and care the weakest, yet for instance Nilsson and Erlandsson (2015) report that fairness and authority both (*r*s = -.44 and .50) have a much stronger association with political orientation than sanctity does (*r* = .29), and Kivikangas and others (2017) report that the left-right orientation has the weakest correlation with the sanctity foundation (*r* = .17, as opposed to *r*s ranging from -.21 to .31 in magnitude for care and authority, respectively). We seek to address whether the differences simply reflect idiosyncratic variation, or whether there are some systematic, theoretically or practically relevant moderators at play. In the following sections, we review the conceptual and methodological issues that guided our identification of potential moderators.

[3] Correlations not provided in the original article. They have been computed on the raw public data and are based on N = 1,209 participants with pairwise complete values: https://dataverse.harvard.edu/dataset.xhtml?persistentId=hdl:1902.1/12658&studyListingIndex=0_775f45d232bb5e430d0024139e25

Issues regarding political orientation and ideology

While ideology has been operationalized in a variety of ways, the most common measure of political orientation or ideology is (given some slight variations) a political self-placement (PSP) item that asks participants where they would place themselves on a single bipolar scale, most often representing a *liberal-conservative* or *left-right* continuum. The simplicity and brevity of a PSP scale has made it easy to implement in a broad range of studies, making it by far the most widely used measure of political orientation or ideology, including the majority of studies linking moral foundations and political orientation or ideology. Although frequently argued to be an adequate measure (Jost, 2006; Jost, Federico, & Napier, 2009; Feldman, 2013; Hibbing, Smith, & Alford, 2014; but see Bauer, Barberá, Ackermann, & Venetz, 2017), its very simplicity may cause confusion as to what exactly it measures. This is not merely a methodological problem, but a theoretical one as well.

Ideology, its dimensionality, and cultural differences

Reviews of (individual-level) political ideology or orientation typically define the term as relating to political beliefs and attitudes about how society should be ordered: advocating vs. resisting social change, and rejecting vs. accepting inequality (e.g., Feldman, 2013; Jost et al., 2009). Personal political orientation has been argued to stem from multiple sources: bottom-up dispositions (the shared psychological structures that allow for individual differences; Federico & Malka, 2018; Hibbing et al., 2014; Jost, Glaser, Kruglanski, & Sulloway, 2003), that interact with top-down influences (communication by the political elites, discursive packaging with other attitudes; Converse, 2000; Zaller, 1992) and form a personal political orientation (Federico & Malka, 2018; Feldman & Johnston, 2014; Jost et al., 2009).

In contrast to the unidimensional model of ideology, an alternative model argues that two distinct dimensions, one encompassing *social* or *cultural* issues, and the other *economic* issues, are needed (e.g., Feldman & Johnston, 2014; Treier & Hillygus, 2009). In this conceptualization, social conservatism refers to resisting (vs. advocating) social change, and economic conservatism to accepting (vs. rejecting) inequality. More generally, the evidence on mass publics around the world suggests that cultural and economic attitudes are typically not aligned on one, but two relatively independent dimensions (e.g., Evans, Heath, & Lalljee, 1996; Feldman, 2013; Malka, Soto, Inzlicht, & Lelkes, 2014). A similar two-dimensional structure reflecting economic and social orientation has also been found in regard to explicit attitudes (Cochrane, 2010), in value research (Schwartz, 1992; see also Piurko, Schwartz, & Davidov, 2011; Thorisdottir, Jost, Liviatan, & Shrout, 2007), and at the level of psychological dispositions (e.g., Carney, Jost, Gosling, & Potter, 2008; Crowson, 2014). Some evidence supports the dual-process model of ideology, which proposes that social and economic conservatism are based on psychological dispositions of authoritarianism (often operationalized with a measure for right-wing authoritarianism, RWA) and social dominance orientation (SDO), respectively (Duckitt & Sibley, 2009).

Federico and Malka (2018) have argued that while cultural and economic issues may have distinct psychological underpinnings, they also may have different associations with the cultural and economic dimensions of political orientation depending on the political context. In the US, ideologies have historically been thought to be unidimensionally structured along a single dimension labelled interchangeably “left-right” or “liberal-conservative”—a convention likely emphasized by increasing political polarization. By contrast, in many West and North European countries, “left” is primarily associated with redistributive policies and “right” with support of low taxes, entrepreneurship, and free markets economic policies, while “liberal” is associated with anti-

discrimination and environmental social policies, and "conservative" with religiousness, traditionalism, and/or nationalism, even if the dimensions are correlated (e.g., Evans et al., 1996). However, the meaning of political labels and the bundle of particular beliefs that comes with a label is determined by the political history of the country, and this meaning changes with demographic changes and current or near-history events (Bauer et al., 2017), such as the influence of communism in Eastern European countries, which pairs conservatism with leftist rather than rightist policies (Aspelund, Lindeman, & Verkasalo, 2013; Piurko et al., 2011; Thorisdottir et al., 2007). In fact, a cross-cultural analysis on the topic suggested that in a worldwide perspective, conservatism is more commonly associated with the left and liberalism with the right (Malka, Lelkes, & Soto, 2019). The issue is further complicated by findings that a large segment of the electorate in fact hold no coherent political beliefs (Zaller, 1992). Instead, ideology as reported by a PSP item will, for the majority of the population, primarily reflect group identification, not views on political issues. The use ideology as a simplifying mechanism to organize broad constellations of beliefs is typically characteristic only for people who are more politically engaged, i.e. politically more involved and knowledgeable (Converse, 1964; Campbell, Converse, Miller, & Stokes, 1960; see Kinder & Kalmoe, 2017, for contemporary evidence).

These considerations give reason to suspect assumptions of Unidimensionality and Universality, as the political differences can be expected to reflect on moral differences. Although Graham and colleagues (2011) reported similar patterns between political orientation and moral foundations in different global regions, this could in part be an artefact of using a unidimensional political axis in contexts where a different model might have been more suitable. We will use moderator analyses to examine those cultural differences that can be quantified in our data, whilst

acknowledging that findings from such analyses are not the same as a proper cross-cultural investigation.

Issues with moral foundations

In contrast to the diversity found in work on ideology, the moral foundations are rather uniformly applied within the literature that builds on MFT. Questions regarding the number and content of the moral foundations (e.g., the liberty foundation, Iyer et al., 2012; the differentiation between physical and emotional harm, Clifford et al., 2015, and Simpson et al., 2016) or the sacralization of the moral foundations (Graham & Haidt, 2012) have not changed research practices. Similarly, criticism from outside and alternative theories (see, e.g., Janoff-Bulman & Carnes, 2013; Gray & Keeney, 2015; Curry, Chesters, & Van Lissa, 2019; see Discussion), have received little attention to date. Given the sparsity of empirical work building on these efforts, there is little possibility to test moderators originating from these criticisms in this meta-analysis. However, we investigate some methodological issues regarding measurement differences and model comparisons.

Issues regarding the moral foundations – political orientation link

Unidimensionality

The implications of the two-dimensional models of political orientation for the associations between more than one dimension and the moral foundations are yet to be explicitly investigated. Most pertinent is a study by Federico and others (Federico et al., 2013; see also Altermatt et al., 2016) who investigated the relationship between moral foundations and the two dimensions identified by the dual-process model of ideology—Right-Wing Authoritarianism and Social Dominance Orientation (Duckitt & Sibley, 2009). Their results show strong positive associations between RWA and the binding foundations, and strong negative associations between SDO and the individualizing foundations. For authority and fairness, the results were as would be expected, given

that RWA and SDO, respectively, were used as external validation criteria in the development of these scales (Graham et al., 2011). However, Federico et al. showed that a similar pattern can be found for all five moral foundations, implying that each one may be distinctively associated with social and economic conservatism. The results from samples gathered in Finland (Kivikangas et al., 2017), and Sweden and Latvia (Dimdins et al., 2016) are consistent with this notion. Besides examining the implications of employing two-dimensional models of political orientation on the associations between political orientation and the moral foundations, we will investigate whether using RWA and SDO in place of single-item self-placement measures of political orientation can shed additional light on how these associations may vary across different models and conceptualizations.

Universality, sampling, and the YourMorals data

If the relationship between moral foundations and political orientation is robust over political cultures and other systematic population differences (Assumption 4), it should not matter where the sample is from or how it was collected. Many Labs 2 study (Klein et al., 2018) tested for some sample and procedural (WEIRD vs. non-WEIRD countries, see Henrich, Heine, & Norenzayan, 2010; lab vs. online collection, number of other measures taken before it) differences, and found little evidence for these factors functioning as moderators. However, as the study relied only on country-level WEIRD-indices, it did not really test for differences between WEIRD and non-WEIRD samples. At the vast majority of the replication sites, participants were local university students (see the project document on sources: https://osf.io/uv4qx/), and university students may in many respects bear little resemblance to non-WEIRD populations, even if they come from a non-WEIRD country (Henrich et al., 2010; see also Inglehart & Welzel, 2010).

Demographical differences have been found to influence the moral foundations endorsement (Graham et al., 2011), and these differences may be more important in some samples than in others. Davis and colleagues (2016) reported that the connections between conservatism and the authority and sanctity foundations are not as strong in the Black population, which is more religious than the White population in the US, but relatively more liberal. Furthermore, they note that the YourMorals data has very few Black respondents, and even fewer of them are religious. They go on to argue that the association between conservatism and the binding foundations may partly be a measurement artefact resulting from biased sampling. These observations give rise to concern because a notable portion of the moral foundations-politics literature has relied on the YourMorals data, which collects the responses of self-selected people who want to learn about morality and who have means and capability to go to the website and fill in the questionnaire in English. This has led to a liberally biased (Graham et al., 2011, Table 2 shows 22,000 self-reported liberals, but only 3,000 moderates and 4,000 conservatives), young, well-educated (and likely not poor), White, and male sample. A distinct possibility is that this sampling systematically biases associations found in the YourMorals sample. Demographic variables such as gender and age are known to be associated with differences in morality—morality for women has been suggested to be less about justice (fairness) and more about care (Gilligan, 1982; see Koleva, Selterman, Kang, & Graham, 2014)—as well as political orientation (a link between age and conservatism is old and commonly known phenomenon; e.g., Truett, 1993), implying that biased sampling on even these very basic demographic variables can distort the results.

In addition to pure demographical differences, self-selection may play a role. A sample of heavily self-selected motivated volunteers—such as the YourMorals data, but perhaps also MTurk respondents, who are allowed to choose the surveys they answer (see Necka, Cacioppo, Norman, &

Cacioppo, 2016)—may be biased due to individual differences underlying the decision to take part in the survey. For instance, it has been long known that liberals are more open to experience, and thus more likely to participate in studies (Dollinger & Leong, 1993; Gerber et al., 2010; Joe, Jones, & Ryder, 1977). Similarly, samples of college students, although not necessarily strictly self-selected (e.g., if participating is part of a mandatory course), are based on a population that may not be only demographically very narrow, but also self-selected in the sense that they have had the opportunity and motivation to pursue higher education.

To account for these potential biases, we will compare results based on the YourMorals sample with results based on other common sample types used in studies independent of the YourMorals data, mostly convenience samples and MTurk samples. In addition, we will attempt to obtain representative samples that have explicitly made an effort to sample also conservative and other less typical groups (e.g., by using a nationally balanced respondent panel; Chang & Krosnick, 2009). If differences between sample types are found, we will further investigate to what extent these can be attributed to demographical differences or to variables related to self-selection, such as interest in politics.

Current Meta-Analysis

We identified four common assumptions in the literature on the associations between the moral foundations and political orientation: (1, Direction) care and fairness are negatively associated, and loyalty, authority, and sanctity are positively associated with conservatism; (2, Magnitude) associations between political orientation and care are the weakest and associations of political orientation with authority and sanctity are the strongest; (3, Unidimensionality) different labels of the political orientation measure are interchangeable in how they are related to moral foundations; and (4, Universality) the associations between moral foundations and political

orientation are similar in different samples and in different cultures. In our analyses, moral foundations are measured by the Moral Foundations Questionnaire (MFQ; see below), and political orientation by political self-placement (PSP) items. Our two goals are to attempt a replication of the established findings, and to investigate variables we identified as potentially moderating the associations. The moderators we identified include: sampling, especially YourMorals compared to representative sample types; demographic variables, especially gender, age, and race; what labels are employed to anchor the measure of political orientation; and respondents' country of origin. Political engagement was also identified as a theoretically interesting moderator, but no study reported it—however, a few studies reported political interest, which reflects a close construct (e.g., Malka et al., 2019, builds political engagement from two variables of which political interest is one). Furthermore, we will examine whether the dual-process model of ideology (RWA & SDO) could help illuminate the psychological tendencies that ostensibly underlie political orientation.

Methodological issues with the moral foundations questionnaire

Besides the theoretically and empirically interesting moderators described above, the variation in measuring the moral foundations warrants moderator analyses on the methodological quality of the data.

Moral foundations—as traits of an individual, as opposed to, e.g., evaluations (Frimer et al., 2013) or differences in moral language (Feinberg & Willer, 2015)—are almost exclusively measured with one primary, validated measure, MFQ (see Graham et al., 2011). It includes two parts: relevance items, which ask what the respondent considers relevant when deciding about right and wrong, assumed to correspond with explicit reasoning about moral arguments, and judgment items, which focus on more concrete and contextualized questions that are assumed to be more related to the intuitive part of morality. Each of the five foundations has three relevance and

judgment items, resulting in a 32-item scale, including also two items that check response quality. An abbreviated 20-item short form consists of the two highest-loading relevance and judgment items per foundation. Most research on MFT is based on 32, 30, or 20 item versions of the MFQ (the latter two lack the quality check items). Some earlier studies utilized an earlier 40-item version, and some studies have used only one subscale or only some of the foundations, resulting in 15 or fewer items (most notably, the Many Labs 2 study used only the relevance items, Klein et al., 2018). As moderators, we are interested how these differences influence the measurement quality (20 vs 30 item MFQ, use of quality check items, using whole MFQ vs just the relevance or judgment items). Although some other measures have been devised (e.g., Moral Foundations Sacredness Scale, Graham & Haidt, 2012; Moral Foundations Vignettes, Clifford, Iyengar, Cabeza, and Sinnott-Armstrong, 2015), they have not been widely adopted and are not included in our analyses.

In the original validation study, the MFQ was reported as having less than optimal psychometric properties regarding internal consistencies, item loadings, and fit indices. Although these issues are mostly out of scope for this meta-analysis, the consistency with which the five foundations have been found to group into two individualizing and three binding foundations has led to doubts on whether the five-factor model is really preferable to a two-factor model. In fact, many independent studies have used only the two higher-order factors (e.g., Napier & Luguri, 2013; Niemi & Young, 2016; Rossen et al., 2015; Yilmaz & Saribay, 2017) instead of each of the five foundations separately. In independent studies, although the five-factor model has outperformed alternative models (Graham et al., 2011; Nilsson & Erlandsson, 2015; Yilmaz, Harma, Bahçekapili, & Cesur, 2016), the differences are not great and the factor analyses have shown fits far below the typical recommendations: for example, CFI = .59 and .68; RMSEA = .10 and .07, respectively (Davis et al., 2016; Nilsson & Erlandsson, 2015). This issue pertains to Assumption 2: if the two-

factor model is consistently superior, there may often be no good reason to consider the moral foundations individually. However, if the two individualizing and three binding foundations do, within these respective groupings, sometimes have clearly pattern-breaking associations to political orientation, then this would present an argument against the two-factor model that would be stronger than methodological arguments based on technical model fitting alone.

Methods

Search procedure

We developed the search procedure and form based on the PRISMA guidelines (Liberati et al., 2009), and followed most of the recommendations by Lakens, Hulgard, and Staaks (2016; we did not preregister our analyses). The forms and the recorded steps are available in the Supplementary Material S1. Disagreements were to be resolved by discussion between the two review authors (JMK and SJ). If no agreement could be reached, it was planned that they were subjected to a vote between the three other authors; however, this was never needed.

For the database search, we used ScienceDirect, Scopus, ProQuest, ISI Web of Science, Ovid PsycArticles, SpringerLink, Taylor and Francis Online, Open Science Framework, SSRN eLibrary, and PubMed[4]. The search term was "Moral foundations questionnaire", as the MFQ was an integral part of the research questions, and it still produced a manageable number of papers in total[5]. The date was limited to a range from 2009 (when the seminal study by Graham et al., 2009, was published) to May 2017. Two of the authors, JMK and SJ, independently carried out the searches from their respective universities that have partly different journal subscriptions. JMK carried out the searches between 13 and 15 June 2017, and SJ between 29 Sep and 2 Oct 2017. We

[4] In addition, we searched Cordis Library, LearnTechLib (formerly EdITLib), and OpenGrey for unpublished studies, but each produced zero hits.

[5] We considered different variations of "political orientation" as search terms, but ultimately decided against them. The terms vary a lot and we could not be sure we did not miss a relevant one, and filtering out those that do not mention some version of the term does not reduce the number by much, because the finding of Graham et al. (2009) is often mentioned even if political orientation is not used in the study.

also checked the articles listed on the MoralFoundations.org publications page on 13 October 2017 for any missed studies. After screening out the duplicates, authors JMK and SJ independently checked the eligibility of the studies according to the eligibility criteria (see below). To test the interrater reliability, we compared the first hundred decisions that both the authors did independently. The interrater agreement was 94 %, Cohen's kappa = 0.866 (see Supplementary Materials S1 sheet Interrater_reliability). The reliability was high enough, so to reduce workload we divided the remaining articles between JMK and SJ to be checked for eligibility separately.

After the eligibility checks, the next step was for us to extract the effect sizes for each MFQ-PSP relationship from the set of eligible articles. The coefficients were not directly available in most of them, so we contacted the authors of the articles that did not report the relationships, and asked them to provide either the calculated correlations or the raw data (the articles that did report the coefficients were asked for the raw data only). We also asked for RWA and SDO variables, any variables related to Schwartz values (which were ultimately obtained from too few studies to be included in the meta-analysis), and demographics. In the same email, we asked whether the authors had any eligible unpublished data we could also include.

Eligibility criteria

We used the following criteria for inclusion of studies (few exceptions are reported later):

1) We include published original studies with using both MFQ and at least one PSP from scientific journals, dissertations, and books, published between 2009 and May 2017, and from unpublished manuscripts obtained during the same time. We set the lower boundary by the publication year of the seminal article by Graham et al., because the MFQ was not fully developed before that. The higher boundary was set during the process to provide a clear cutoff.

2) Seeing that YourMorals data has been used in multiple studies, we avoid including the same participants multiple times by including the YourMorals data only once. This also applies to other popular data sources, if found. All (further) samples that directly use the YourMorals data, or from ProjectImplicit.org data which overlaps with it (according to B. Nosek, personal communication, April 18, 2016), are excluded. If a study uses both YourMorals data and an independent sample, only the independent may be used. This criterion was later applied to other databases that we found used in the literature (see Study selection and data summary, below, for more detail).

3) To assess political orientation, studies to be included must have used at least one PSP item on a) at least a 5-point scale, b) with instructions referring to either political ideology or orientation, c) and anchors labeled either liberal-conservative or left-right (with or without possible qualifiers such as social and economic conservatism). The vast majority of the relevant studies used a 7-point scale, and to ensure sufficient variation, we decided to include only studies with 5-point scale or wider. Criteria 3b and 3c were decided by the fact that, based on preliminary searches[6], these were the most widely used measures. We rejected the few found studies with issue-based measures and measures that depended on the US-centric party identification. Studies using RWA and SDO were not specifically targeted, but we asked for those measures in addition to raw data when contacting the authors.

4) To assess MFs, studies to be included must have used the MFQ (either 20-item or 30-item version). Other measures or versions (such as the older 41-item version) were rejected on the basis that their validity is not clear, and they were used too rarely to use MFQ version as another moderator in the analyses. Studies that did not employ at least four MFQ factors were also excluded.

[6] Preliminary non-documented searches (results reported in Kivikangas, Lönnqvist, Ravaja, 2016) were done in winter 2015-2016, and for instance the number of articles using other measures of moral foundations than the MFQ were assessed based on them.

5) Only studies using moral endorsement of the respondents themselves, naturally occurring, are included. Some studies have administered the MFQ after an experimental manipulation or instructed the respondents to answer as they feel someone else would answer, which both artificially influence the main measure. Such studies were excluded on the basis that they are not indicators of how moral foundations and political orientation are generally associated.

The study-level effect sizes included are Pearson's correlation coefficients between the PSP measure(s) and the five individual MFQ factors representing the five moral foundations. We considered collecting secondary effect sizes such as regression weights and partial correlations, but ultimately, we rejected them, as we could calculate correlation coefficients with the raw data for many of them, and the remaining few instances were too few to be analyzed separately and using them together with bivariate correlations is not recommended (Aloe, 2015).

As a practical limitation, we did not specifically search for studies published in other languages than English. However, we include a considerable number of studies that were conducted in other languages but reported in English, as the topic is internationally interesting. In addition, the search term itself (see below) was a reference to the name of the MFQ, so non-English articles also showed up in the searches when they refer to the questionnaire. We found Russian, French, Italian, and Latvian records within the search results, and contacted the authors in order to find out whether the studies were eligible. Those articles whose corresponding authors answered, found eligible, and provided the required information were included.

Study selection and data summary

A flowchart on the selection of studies is in Figure 1. The literature search produced a set of 939 records, and 112 articles were found by other means. The records and articles were screened for duplicates, obvious non-eligibility cases and lack of availability, resulting in 389 full-text articles

which were assessed for eligibility. If something was unclear, the corresponding author was contacted. We also requested authors for access to the raw datafile in order to make more detailed analyses. Ultimately, we contacted around 99 authors, of which 24 authors of otherwise eligible studies either could not be contacted at all, responded at first but later could not be contacted again, or declined to share data. When the raw data was not shared, the published effect sizes were nevertheless used when available. The final inclusion list has 53 journal articles, 7 theses (including both doctoral and master's), 4 unpublished datasets (of which two were published only as posters) and 3 databases, containing in total 89 samples. All samples provided from five (between one PSP item and five MFQ factors) to fifteen (three PSP items and five MFQ factors) effect sizes. The total number of effect sizes was 605. A list of included samples, with information on their publication statuses, samples sizes, relevant moderators, and the effect sizes can be found in Supplemental S3 Tables S3T1 and S3T2.

The three databases used were the YourMorals dataset by the original authors of the MFT, American National Election Studies (ANES), and New Zealand Attitudes and Values Survey (NZAVS), each of which were used in several articles but were included as samples in the final *k* only once to avoid overlap. The YourMorals data, used by 29 records in our data selection process, was requested from Jesse Graham and obtained on 27 May 2016 with total of 218,330 participants. Three articles (Clifford, 2017; Jones, 2011; Koleva et al., 2014) reported using ANES data, which was downloaded from https://electionstudies.org/data-center, limited to respondents participating both in 2008 Wave 6 Panel data in June (which employed a PSP measure) and Wave 7 off-panel data in July (the only ANES questionnaire which employed MFQ)[7], a total of 1149 participants. Four articles (Bulbulia, Osborne, & Sibley, 2013; Davies, Sibley, & Liu, 2014; Malka et al., 2016;

[7] Although the ANES data was collected in 2008, before our start date, we included it because of the general lack of representative samples, after ensuring that they did use the eligible MFQ version and not the old one.

Milojev et al., 2014) reported using NZAVS, and we obtained this dataset from Chris Sibley on 8 May 2016, with total of 23,742 participants, although the majority (more than 18,000) of the cases were discarded due to not including both focal measures.

The total *n* for MFQ-PSP correlations from the 89 samples was 226,674, of which independent (i.e., non-YourMorals) studies ($k = 88$) $n = 33{,}804$, and YourMorals ($k = 1$) $n = 192{,}870$. The analyses that are conducted on individual-level data included 48 samples that—excluding the YourMorals dataset—had a total $n = 22{,}028$. Of the total number of independent samples, convenience samples (students, snowball samples, and online samples with unspecified sampling strategy) comprised the majority with 57 samples, and most of the others were collected using MTurk ($k = 27$). Only four samples were reported using a strategy to obtain a sample representative of the general population (three online panels, and ANES using computerized phone interviews). More than half of the samples ($k = 50$) employed a PSP measure with a liberal-conservative self-placement item, and the rest used—either exclusively or in combination with another measure—a left-right item ($k = 24$; some of the samples used both lib-cons and left-right), or two separate social orientation and economic orientation items ($k = 21$; a few samples had another [e.g., foreign issues] conservatism item in addition to social and economic orientation, which we excluded).

The included literature is heavily focused on the US, with 47 of the samples. Seventeen samples were collected in Europe (six in Italy, five in Sweden, three in Latvia, two in Finland, and one in the UK), and the rest in individual countries around the world (notably five from Turkey), international population, or unspecified countries of origin.

Five studies reported using the same 16-item measure for SDO (Pratto et al., 1994; Sidanius & Pratto, 1999). For RWA, Altemeyer's (2006) original 22-item measure, or different short forms

of that instrument were used (Federico, Hunt, & Ergun, 2009; McFarland, 2010; Sibley & Duckitt, 2009; Zakrisson, 2005). For the two remaining studies (Hofmann et al., 2014; NZAVS), we did not find information what were the exact measures they used.

Excluded studies

Apart from the explicit inclusion criteria, the main exclusion reason was key values being not reported in the article or reported in wrong units, in combination with our inability to reach the authors for questions or data (notably, e.g., Davis et al., 2016). Other major reasons were an experimental manipulation included before the respondents answered the self-reports (e.g., Bassett, Van Tongeren, Green, Sonntag, & Kilpatrick, 2015; Wright & Baril, 2011; Napier & Luguri, 2013; Cornwell & Higgins, 2014); using a too coarse measure (less than 5 points), or used US party affiliations as a measure for political orientation (e.g., Low & Wui, 2016); using only some of the foundations (e.g., Royzman, Atanasov, Landy, Parks, & Gepty, 2014); or using only the relevance or judgment items (e.g., Many Labs 2; Klein et al., 2018).

Assessing the quality of the data

The typical quality assessment methods are designed for clinical studies and especially randomly controlled trials (e.g., Moher, Jadad, & Tugwell, 1996; Sanderson, Tatt, & Higgins, 2007), which studies relevant to our focus did not use. A large portion of the relevant studies did not use an experimental design at all, and those that did, used irrelevant manipulations. However, as data quality issues are central to our focus, we test the sampling method and measure differences discussed in the introduction as moderators. In addition, we test the differences between studies using short or full forms of the MFQ, and studies that use or do not use the two MFQ quality check items (“Whether or not someone was good at math” and “It is better to do good than to do bad”), designed to identify responders who are not paying attention or responding seriously. Regarding

other indicators of data quality, preregistration procedures have been proposed (e.g., Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 2012), but the relevant literature did not include any preregistered studies. However, our confidence in the results and conclusions is increased by the fact that the relationship between moral foundations and political orientation has not been the main focus of any of the independent studies. This means that obtaining any particular value for this relationship cannot, at least directly, have been an undisclosed selection criterion for the independent studies.

Further, we use three methods to detect publication bias. Funnel plots are plotted using standard error as the *y*-axis, as recommended by Sterne and Egger (2001); the other often recommended choice of precision (1/se) emphasizes large-sample studies, which we wanted to avoid due to the size difference between the YourMorals dataset and the independent studies. In the absence of publication bias, it is expected that studies form an inverted and symmetric funnel shape, that is, small studies (at the bottom of the graph) are expected to be more spread around the pooled effect size, compared to large studies (plotted at the top of the graph) that are expected to be closer to the combined effect size. If the funnel plot is asymmetric, especially at the bottom, it suggests the presence of a publication bias. Egger's regression test is a popular method to use in conjunction with funnel plots, but simulation studies have repeatedly shown that this method leads to very high Type I error rates under many circumstances (Kromrey & Rendina-Gobioff, 2006; Macaskill, Walter, & Irwig, 2001; Sterne, Gavaghan, & Egger, 2000). Instead, we applied an improved version of the *p*-uniform method (van Assen, van Aert, & Wicherts, 2015). This method works under the assumption that observed significant *p*-values (i.e., *p*-values smaller than 0.05), conditional to the true effect size value, follow a uniform distribution. If the distribution of the significant p-values is not uniform (typically massing just below $p = 0.05$), it suggests the presence of a publication bias. In

the original version of the *p*-uniform method, only the distribution of the significant *p*-values is inspected, while in the improved version[8], called *p*-uniform*, non-significant effect sizes can be also included in the analyses, making this estimator more efficient (van Aert, 2018).

Finally, we did a number of sensitivity analyses. Most prominently, we used the selection method of Vevea and Woods (2005), which is based on the Vevea and Hedges (1996) method but is more suitable when there is a small number of studies within a category. The method compares the unadjusted estimate of the overall effect size and the between-studies variance under a random-effects model to the adjusted estimate of the overall effect size and the between-studies variance using the selection method with the pre-specified weights modeling publication bias. If the adjusted and unadjusted effect sizes are substantially different, it indicates publication bias. The selection method was applied to each MFQ factor-PSP measure combination separately, but as this method is not adapted to the situation where multiple effect sizes are reported within studies (such as social and economic conservatism), we randomly selected one effect size per study. We specified the *p*-value intervals and weights, referring to the probability that a study with a specific *p*-value is published, using the recommended values (Vevea & Woods, 2005).

Analysis strategy

Due to the exploratory nature of these analyses and to avoid data dredging, we do not report significance tests for the moderator analyses except in pre-planned cases, and rather compare the effect sizes. Two correlations have been considered as substantially different if $\Delta r > .10$. This cutoff was selected using empirical criteria. By looking at the distribution of the correlations, we found that the difference between the first quartile and the median, and between the median and the third quartile is about .07 units across all MFQ factors. In addition, after the sensitivity analyses regarding the influence of the use of different subsets of the MFQ (30 vs 20 items, use of attention check

[8] Used via the Shiny app at https://rvanaert.shinyapps.io/p-uniformstar/

items, use of only the relevance or judgment half) and the selection method by Vevea & Woods (2005), we found that the maximum effect these methodological differences had was close to $|r| = .07$. Taking all these factors into account and adding a small marginal, a minimal relevant difference of $\Delta r > .10$ was considered reasonable. A similar cutoff was determined for the regression coefficients from the mixed model analyses. An unstandardized effect of $b = .143$ corresponds to a change of the MFQ score by one, a notable increase on a six-point scale, when moving from the liberal to the conservative end on the PSP scale. Considering the standard deviations of each MFQ factor across all studies, a minimal relevant difference of $b > 0.8$ was decided.

The only multiple comparisons we do are pre-planned, and within an individual univariate analysis are corrected by the Tukey method. We employ alpha = .01 instead of the more conventional alpha = .05 when significance tests are used, as inferences regarding these are based on five univariate analyses (one per MFQ factor), while still reporting values where $.05 > p > .01$. All conducted analyses, including those ultimately not used in this manuscript, can be found in the Supplementary Materials S2.

We recognized that regardless of the huge n, using the YourMorals dataset as a group with $k = 1$ in the study-level analyses would lead to severe lack of power to test the differences between categories. Thus, instead of using misleading estimates, we calculated CIs directly from the raw individual-level data.

We start from the simple overall analyses corresponding to the regular level of scrutiny in the literature, before proceeding to the more sophisticated analyses and models that use more information and are more realistic. We expect that the simple analyses at least partly replicate the

original results that have supported the target assumptions and we use the more detailed analyses to find out whether that support is misleading when the issue is scrutinized in more detail.

The most basic and the most comparable analysis to those conducted in the literature is a set of five univariate meta-analyses, one for each association between political orientation and a moral foundation. In this step, we use the average of all PSP measures each study has used, following the averaging practice in the literature and the assumption that there are no systematic differences between the item anchors/labels and no consideration of dependencies between individual effect sizes. This analysis includes all the studies, while the more sophisticated analyses in the later steps require information that was not available in every study.

For the rest of the study-level analyses we use meta-analytic three-level models (recommended by Cheung, 2014; and Van den Noortgate, López-López, Marín-Martínez, & Sánchez-Meca, 2013). While a basic meta-analysis only models pre-estimated sampling variance based on sample size (corresponding to level 1 in the three-level model) and the between-studies variance (corresponding to level 3), a three-level meta-analysis also models the within-study variability between effects that originate from the same study. Within studies, the MFQ-PSP correlations with several PSP measures generate dependencies among them (e.g., between social and economic conservatism measures when both correlated against the same MFQ factors). By applying a three-level model, this dependency among correlations is taken into account, without the need to use to common alternatives such as averaging over or selecting only one among many effect sizes, which would have probably biased the results and reduced the statistical power (Moeyaert et al., 2017).

We also use alternative methods to take dependencies into account. Robust Variance Estimation (RVE) is a method that corrects the standard errors (by unshrinking them) to avoid

getting false positive results. We reran the analyses twice, using the RVE methodology and common two-level meta-analyses, and using the RVE methodology together with the application of three-level models, following the recent recommendations for this combination (Tipton, Pustejovsky, & Ahmadi, 2019). However, RVE is not reliable when the number of studies within a category of the moderator variable is smaller than 5, and because in no other case the differences between these methodologies were above $\Delta r = .05$, we use the three-level models without RVE as our default.

For additional sensitivity analysis, we note that in some analyses we use studies that had singular uncommonly anomalous outlier values among otherwise common correlations following robust patterns. Due to suspicion that these correlations are typos rather than reflect real phenomena we ran the analyses without these values for comparison purposes (retaining the other correlations from that study). These changes were inconsequential in the end as well.

The analyses were run using R (R Core Team, 2015), with the *metafor* version 2.0-0 (basic univariate analyses without moderators; Viechtbauer, 2010), *metaSEM* version 1.2.0 (three-level meta-analysis; Cheung, 2014), *weightr* version 1.0.1 (RVE methods; Coburn & Vevea, 2016), and *lmerTest* version 3.0-1 (mixed models, see below; Kuznetsova, Brockhoff, & Christensen, 2017) packages.

Individual-level analyses

Main moderator analyses attempted to explore the effects of demographic variables using meta-regression analyses where study-level demographic variables were included in the regression model (e.g., % of men, % of Blacks, % of Christians), but as very few studies reported demographic information, the power was very low. Therefore, a series of multilevel models were performed using the individual-level data where the demographic information was reported for every participant, with participants (level 1) nested within studies (level 2). Individual-level data have greater power,

as they utilize the whole data instead of a single descriptive per study, and they are able to detect patterns that may remain hidden in the study-level analyses (Riley, Lambert, & Abo-Zaid, 2010). For the data where those items were available, we excluded participants who had answered the question “math” with the “somewhat relevant” to “extremely relevant” (upper half of the scale), or the question “good or bad” with “strongly disagree” to “slightly disagree” (lower half of the scale).

Before performing the analysis, all PSP measures were rescaled to the range from 0 to 6 (the most common range in the data). The aim was primarily to find out whether the differences between sample type specific PSP slopes are smaller compared to each other, which would indicate the extent that the differences between sample types are attributable to the moderator. Because each of the moderator analyses used a different subset of the data, we checked whether the non-moderated predictions using these subsets were reasonably equivalent to the non-moderated prediction using the whole dataset. No notable differences were found for binding foundations (all differences < 0.014, relative change < 4%). For care and fairness, both absolute and relative differences between data subsets were slightly larger (absolute difference 0.027 and 0.022, 16.4 % and 12.0 %, respectively; largest changes in the subsets for education and religion in the care model, and for education in the fairness model).

The first step was to replicate the results found in the study-level analyses with the individual-level data in the null models. Second step consisted in incorporating one demographic moderator variable in the model, and its interaction with sample type (hence controlling for the difference distribution of the demographic variable across types of samples) and the interaction between a PSP measure and the demographic variable (hence controlling for differences in the relationship between the PSP measure and MFQ factors across the values of the demographic variable). Separate meta-regressions were carried out for each demographic variable with the aim of

maximizing the power of the analyses, because there was little overlap between studies' inclusion of the variables.

For specific research questions, models with simultaneous moderators were performed, following recent recommendations on meta-regression analyses (Tipton, Pustejovsky, & Ahmadi, 2019). However, due to the poor overlap of moderator variables in the primary studies, these analyses are run on very restricted datasets and are consequently less reliable.

Moderators

Sample type

We employ a moderator with four categories, each corresponding to a **sample type**: YourMorals, convenience, MTurk, representative. The sample type moderator was coded as '*YM'* for the YourMorals data; '*convenience'* when the original article reported that respondents were recruited in such ways as a convenience sample explicitly from students, more vaguely at the university, or using an online survey without specifying representative sampling; '*MTurk'* when the article reported using the Amazon Mechanical Turk respondents or similar services; and '*representative'* when the article reported using a strategy to ensure at representativeness. A representative sample that successfully draws from all social strata is much more likely to provide unbiased estimates. The sample type was reported unambiguously in all original articles.

Demographics and political engagement or interest

Demographical variables for **age** and **gender** (in study-level analyses, % of male in sample) were found in majority of the studies. Other moderators are based on fewer samples. Self-reported scales of **education** were converted to a five-point scale: Less than high school, High school graduate, Some college, Undergraduate degree, and Graduate degree. **Race** was coded with four categories: '*White*', '*Black*', '*Hispanic*', and '*Other*'. **Religion** was coded with three categories:

'*Christian*', '*No religion*', and '*Other*'. In study-level analyses, both race and religion were coded as percentages of the total sample, and to avoid the dependency between the categories, the category 'Other' was left out of the analyses. While there was no direct measure for political engagement in any of the included studies, a proxy measure, **political interest**, could be constructed on a three-point scale, from "Not at all interested in politics" to "Very interested in politics".

Political orientation measures

The relevant differences of a single-item self-placement scale lie in the instructions and anchors—i.e., the orientation label used. A moderator with four categories was formed. The **orientation label** moderator was coded as '*conservatism*' when the measure used "liberal" and "conservative" or "liberalism" and "conservatism" as anchors, with or without labels on the intermediate steps; and '*left-right*' when a measure with "left" and "right" (with or without "-wing") as anchors was used, with or without intermediate labels. Measures labeled with phrases like "social conservatism" and "economic conservatism" (or more rarely, social and economic right-orientation) were coded as '*social orientation*' and '*economic orientation'*. In the rare cases where larger number indicated higher liberalism instead of conservatism, the scales were reversed. Sometimes also "general conservatism" or "foreign policy" were used, but those were ignored due to their rarity when we obtained the separate items and are present only as part of the aggregates if we did not obtain the separate items. When unclear, the exact wording of the anchor labels was asked from the authors of the original article. This revealed a couple of deviations from the typical formulations, such as studies combining liberal-conservative and left-right anchors in the same item when this was not explicitly reported in the article (see Supplemental Table S3T1 for details).

For an alternative operationalization of social and economic conservatism, RWA and SDO were employed. We calculated the correlations between MFQ factors and both RWA and SDO, using those correlations as effect sizes in the same way as the PSP items are used in other analyses.

Political culture

Due to a lack of samples for most of the countries, we chose to use the study-level analyses to test the interaction by comparing two major **regions** the most of the studies were from, *the United States* and *Europe*, as groups with a 2x2 interaction with the PSP measures. In the individual-level analyses, we could test the effects of specific countries, although the more precise interaction analysis could only be used for those studies that used both relevant measures.

Rejected moderators

We also considered many other moderators that were rejected them from the current study. Some studies have suggested that making moral or political judgments in a stressful or loaded situation influences these judgments, albeit their conclusions on the direction of the effect differ (Eidelman, Crandall, Goodman, & Blanchar, 2012; Wright & Baril, 2011). As there were only very few studies investigating these effects, we decided to restrict outside our analyses the measures or constructs such as the Big Five personality traits (Gerber et al., 2010), Machiavellianism or dark triad traits (Niemi & Young, 2013), Schwartz basic values and political values (Dimdins et al., 2016), emotion regulation (Feinberg, Antonenko, Willer, Horberg, & John, 2014; Feinberg, Willer, Antonenko, & John, 2012; Graham et al., 2012), or framing effects (Day et al., 2014).

In addition to the differences in anchor labels, the use of the key instruments was not completely uniform across the studies. For instance, some studies asked about political “ideology” while others about “orientation”. As the distinction between these is not always made even by the scholars of political psychology, we assumed that this distinction is not relevant in the sense that

these choices would introduce systematic biases. Similarly, we ignored whether 5-, 7-, or 9-point or some other scale was used for PSP, and whether and how the PSP scale midpoint was labeled ("independent or unsure", "moderate", "middle-of-the-road" etc.).

Results

Main analyses

Although we did not ask for raw MFQ item scores from the authors, we could calculate some basic descriptive statistics for the subset of samples from we had raw data ($k_{max} = 48$). The descriptives for these samples are shown in Table 1.

The mean MFQ factor scores are highest for care and fairness, and lowest for sanctity, but standard deviations are highest for sanctity, indicating least agreement. The relationships of these statistics between different foundations are also aligned with those reported by Graham and others (2011), although the absolute means are higher for binding foundations, especially sanctity (M = 2.47, vs. M = 1.54 in Graham et al.). All MFQ-PSP correlations range to both sides of zero, indicating at least occasional disagreement on the Direction and Magnitude assumptions.

The internal consistency ranges from .612 (care) to .824 (sanctity), indicating that the each MFQ subscale has substantial portion of variance attributable to the individual items instead of the general factor. However, the internal consistency over these 36 studies is not substantially different from the original study by Graham (2011), where α ranged from .65 (fairness) to .84 (sanctity), so the relatively low alphas appears to be a feature of the instrument rather than the data.

Unmoderated effects and heterogeneity

Using all samples and the conventional PSP measure averages, the random-effects univariate estimates (Table 2) are in the same direction (negative for care and fairness, positive for loyalty, authority, and sanctity), and similar in relative magnitudes (the largest effect sizes for authority and

sanctity, smallest for care) as suggested by the Direction and Magnitude assumptions. With individual-level data (k = 48), the baseline unmoderated mixed models show that PSP predicts care the least but still significantly, and sanctity the most. The results follow those found with the study-level analyses above, and adhere to the Direction and Magnitude assumptions in the same pattern (bs = -0.167 (care)[9], -0.186 (fairness), 0.250 (loyalty), 0.308 (authority), and 0.392 (sanctity), smallest $t(232{,}800) = 163.3$, all ps < .0001). Heterogeneity, reported in Table 2, is discussed in detail in the Supplemental S3, but in brief, tau reflects the dispersion of effect sizes between studies in terms of the original scale, i.e. here, correlation coefficients. As the estimate of the standard deviation of the true effect sizes (assuming normality), tau ranging between .088 and .113 indicates a rather small amount of heterogeneity between samples in absolute terms, but compared to the effect sizes (ranging from more than half the care effect size, to about a quarter of the authority effect size) the heterogeneity could be considered moderate. Compared to the Many Labs 2 study (Klein et al., 2018), which used a unified study design across samples and reported I^2 = 64% and tau 0.09 (in Cohen's d, corresponding roughly 0.05 in correlation coefficients), the heterogeneity across the independent samples is somewhat higher.

Sample type differences

Three-level models, one per MFQ factor, show that with a single exception (representative vs. convenience, in comparisons for authority), the estimate differences are consistent (see Table 3). The estimates based on the representative sample types suggest the weakest associations, while convenience samples indicate somewhat stronger, MTurk samples still stronger, and YourMorals samples clearly the strongest associations. The differences between the representative, convenience, and MTurk sample types are small (Δr < .1), and in (Tukey-corrected) pairwise comparisons not

[9] Note that all regression coefficients for individual-level analyses are unstandardized and referring to changes in MFQ scores (on a 6-point scale) relative to changes in PSP (standardized to the most common, 7-point scale). Thus the effect sizes are not directly comparable to the study-level correlation coefficients.

significant (except for authority factpr, convenience vs. MTurk, Δr = -.067, z = -2.875, p = .018; all other ps > .05). Compared to the Many Labs 2 (Klein et al., 2018) estimates over the higher-order factors, the representative sample shows roughly equal effects for individualizing foundations (averaged r = -.13 in ML2), but the effects for binding foundations are still clearly larger in all sample types (compared to averaged r = .14 in ML2). On the other hand, the very high correlations from the YourMorals sample are, in case of the weakest associations, more than double compared to the estimates by representative samples and ranging from Δr = .143 (authority) to Δr = .253 (fairness), suggesting significant inflation.

Using alternative models with RVE alone, and RVE within three-level models, the estimates changed slightly and the standard errors were larger for convenience. SEs were smaller for MTurk and representative sample types, making some differences between individual sample types statistically significant, but the effect sizes of these differences remained $\Delta r < .1$.

Concerning relationships between foundations, all the sample types follow the pattern explicated by the Direction assumption, and almost all follow the Magnitude assumption, with lowest estimates for care, and highest for sanctity. This suggests that the inflated nature of the YourMorals data is uniform across foundations and does not involve an additional systematic bias.

The overall results of the MFQ-PSP associations are replicated in the individual-level analyses. In the models with sample types separated, the interactions between PSP and sample type confirmed that the MFQ-PSP associations are higher in the YourMorals data—$0.12 < \Delta b < 0.15$ for care, fairness, and sanctity, but only Δb = 0.08 for loyalty and authority, compared to representative sample type—while the differences between the independent sample types were clearly smaller (Δb < 0.05, except for sanctity, for which Δb = 0.07). These analyses are used as the null models when comparing to further moderator models in individual-level analyses below.

Moderating effect of demographics and political interest on sample type differences

Potential causes for the difference between the YourMorals and independent data include differences in demographics and political engagement (proxied by political interest in our analyses). Table 4 presents the means and number of samples for demographics and political interest variables. The samples differ in many aspects, but the main difference of interest is between YourMorals and the independent sample types. The YourMorals sample shows a higher portion of males than the other sample types, is slightly more educated, has less Black respondents, and shows higher political interest (cf. Graham et al., 2011). Age, Hispanic %, Christian %, and no religion % are not consistently different from all other sample types. The percentage of White respondents shows no difference.

We ran study-level moderator analyses with these moderators, but because they are in many cases based on very few studies (except for age and gender) they resulted in very low power and uncertain conclusions, and because the individual-level analyses offer a much more precise view on the moderator effects (from which the study-level analyses do not considerably differ), we report the study-level moderator analyses only in the supplementary materials (see Supplementary S2 for details). As a summary, the only study-level moderator consistently explaining differences between the YourMorals sample and independent sample types was the percentage of Black respondents.

While inclusion of the moderators age, gender, and education in the model decreased the care and fairness associations with conservatism in many cases near to zero in the independent sample types, and the inclusion of religion decreased all associations (including the binding foundations-PSP associations and the associations in the YourMorals sample), none of these moderators explained the difference between the YourMorals sample and independent sample types. Compared to the representative sample type, the PSP main effect for the YourMorals sample is in

almost all MFQ-PSP cases $\Delta b > .1$, and the PSP × moderator interaction effects were small ($\Delta b < .1$). For instance, theoretically interesting gender effect was only $\Delta b = .063$ between at the strongest (for care, the effect stronger for males).

When political interest was included, the differences between YourMorals and the independent sample types disappeared in the models for care and fairness, so the MFQ-PSP association was revealed to be strongly dependent on the moderator. We reran the models with 3-way interactions (Table 5). The main and interaction effects for PSP show that the care-PSP and fairness-PSP associations were heavily dependent on political interest in the YourMorals sample, while the binding-PSP associations were almost independent of political interest. However, due to lack of data that included the political interest variable, the analysis included only five studies.

Modeling the influence of race with the three-way PSP × sample type × race interaction (Table 6), these analyses revealed large differences for all foundations. In all sample types, all the MFQ-PSP associations are the smallest for Black respondents (in the only representative sample here [ANES], the largest $b = 0.102$, for sanctity, compared to 0.337 in the YourMorals sample; for White respondents $b = 0.174$ in representative and 0.410 in YourMorals sample). To a smaller extent, the effects are also smaller in Hispanics and the "other" category, compared to White. In the YourMorals sample, while the differences exist, they are much smaller, and do not explain the differences between the YourMorals and independent samples. Notably, the analyses on race were restricted almost solely to samples from the US, as the studies from other countries did not report race.

We also ran an exploratory model where the sample type was combined with the most potential moderators, political interest and race. The number of participants per independent sample type were too small to interpret the estimates in detail, but the results based on the YourMorals

sample suggest that due to these effects together the individualizing-PSP effects disappear completely (dropped from b = -.22 and -.23—care and fairness, respectively, for White respondents who are "very interested" in politics—to zero—for Black respondents "not interested" in politics), and that political interest was a stronger influence, showing almost double total effect compared to race. Although political interest had only negligible effect on binding-PSP associations in White and the 'other' categories, the authority-PSP and sanctity-PSP associations were influenced by political interest more in the Hispanic (total Δbs = 0.095) and Black (total Δbs = 0.135 and 0.161 for authority and sanctity, respectively) groups.

Political orientation measure differences

Using study-level meta-analyses, we investigated whether the MFQ-PSP associations differ when using different PSP measure. Table 7 shows that the differences between conservatism and right-orientation were very small (Δrs ≤ .06), except less so in case of sanctity, Δr = .13, p < .001). Notably, conservatism is less associated with care and fairness, and more associated with loyalty, authority and sanctity, while such pattern is not true for right-orientation. The individual-level analyses (k = 40) likewise indicate that the differences between conservatism and right-orientation were rather negligible (largest Δb = 0.043, for sanctity). Notably, because the labeling of the PSP measure in YourMorals sample combined conservatism and right-orientation, that sample could not be included in the analysis.

We repeated the analyses for comparing social and economic orientation, for which all data was from the US, except for the YourMorals sample (that had unambiguous measures for social and economic orientation), which was international. Table 8 shows that the study-level difference between PSP measures is significant in case of binding foundations: ranging from Δr = .1 in case of loyalty, to Δr = .2 in case of sanctity. The direction of effects repeats the pattern in the previous

analysis, with social orientation, compared to economic orientation, being less associated with care and fairness, but more associated with loyalty, authority, and sanctity. In regard to the research questions, while all the effects still follow the Direction assumption, they violate the Unidimensionality assumption, but also the Magnitude assumption, with MFQ-economic orientation associations that are about equal in size. The results of the study-level analyses can be seen in Figures 2 thru 11.

The individual-level analyses with a smaller sample size (k = 14) replicate these results, in that the differences between social and economic orientation labels were much larger than between conservatism and right-orientation. Economic orientation was not related to care and fairness particularly more than the social orientation (b = -0.166 vs. -0.109, for care, and b = -0.167 vs. -0.139, respectively, for fairness), but vice versa was true for binding foundations (b = 0.153 vs. 0.240 (loyalty), 0.182 vs. 0.297 (authority), and 0.194 vs. 0.418 (sanctity); all ps < .0001).

We ran a further three-way interaction analysis, which added the sample type moderator to the above model, to examine the PSP label differences across YourMorals and independent sample types (Table 9). Representative sample type could not be used, because none of the samples used social/economic orientation as their PSP measure. The analysis revealed that along with YourMorals sample showing the largest effect sizes, it also shows largest differences between the political dimensions, indicating that the effect size inflation emphasizes bidimensionality.

In addition, we investigated whether the bidimensionality was dependent on political interest. We ran the moderator models separately with YourMorals sample and social/economic conservatism, and with the conservatism/right-orientation sample set with two other studies that had employed the political interest measure. In both analyses the effect was nonexistent (largest b = 0.016); that is, political interest did not appear to vary with bidimensionality.

Differences in political orientation measures by region and country

To investigate the differences in PSP measures across different political cultures, we formed regional groups for effect sizes originating from the US and Europe, which were the only groups large enough to be analyzed. We could only compare PSP measures using "conservatism" and "right-orientation" as labels, as social and economic orientation measures were only used in the US. We used another set of study-level models with two moderators, PSP measure and region, and their interaction. In Europe, conservatism had weaker association with care and fairness, and stronger association with sanctity (Δr = .11 for care, to .19 for sanctity), compared to right-orientation (k_{cons} = 6, k_{right} = 14)—replicating the pattern found in the previous analyses. In the US, all MFQ factors had very close to equal relationships to both PSP measures (all corrected ps > .6), but the analysis included only one study for right-orientation (while k_{cons} = 29). When comparing the regions to each other, in the US right-orientation was less negatively related to fairness, and more positively to loyalty and sanctity than Europe (Δr = .10 for fairness, to .18 for sanctity). For conservatism the tendency was that in the US it was more negatively related to care and fairness, and more positively to loyalty compared to Europe, although the differences remained just below Δr = .10. In other words, for the binding foundations both conservatism and right-orientation showed a stronger association in the US than Europe, but for the individualizing foundations the association was stronger with right-orientation in Europe, but stronger with conservatism in the US. Together, these PSP labels followed the bidimensional pattern in Europe, but not in the US.

These differences were based partially on very few studies, and the results for Europe were particularly influenced by three samples from Latvia—an Eastern European country with different political culture and history. When those studies were removed (k_{cons} dropping from 6 to 3), the difference between PSP measures in relation to care decreased notably, suggesting that this

difference may be a Latvian or Eastern European rather than European. Some other differences diminished somewhat as well but much less. When using the alternative RVE models, the results changed notably for comparisons involving right-orientation and the US, but as the number of studies with that combination was one, these analyses cannot be considered reliable.

In addition to regional analyses, the individual-level raw data allowed us to compare the associations with conservatism and right-orientation labels in individual countries directly (Table 10); this reduced the analysis to only five studies, a single study per country (n = 6710), recommending caution in interpretation. Most notably, the US is the only country for which Unidimensionality appears to hold; for all other countries there are non-trivial differences between conservatism and right-orientation measures in binding (but not care and fairness) foundations, most strongly in sanctity. The differences follow the familiar pattern: right-orientation has stronger associations with individualizing foundations and conservatism with binding foundations. The effect sizes are in total also largest for the US and New Zealand. Latvia is the only country that breaks against the assumption of Direction, with a positive care-conservatism, and negative right-authority and right-sanctity associations. The total differences between the two measures are also largest for Latvia but at the same time the total effect sizes are the smallest. Magnitude assumption holds generally for conservatism, but not for right-orientation. Finally, Finland and Sweden as culturally similar Nordic countries have similar associations.

As reliability of a scale can influence results (Westfall & Yarkoni, 2016), and there has been some concerns about the reliability of the MFQ, especially when translated, we checked whether these findings are affected by particularly unreliable studies. We found that Cronbach alphas were higher in the US and New Zealand samples (α = .808 and .900, respectively, for sanctity; see the whole table in Supplement S3T8) and lower in the Latvian sample, particularly for authority (α

= .361), but this foundation subscale is not the only one with a stark difference, and it follows the findings for subscales with better alphas (α = .789 for sanctity).

Political orientation operationalized as RWA and SDO

We employed RWA and SDO as an alternative operationalization of political orientation. Table 11 shows that the MFQ associations to RWA and SDO replicate the pattern found between the PSP operationalizations, in that RWA (like social orientation in the US, and conservatism in Europe) is more strongly and positively associated with loyalty, authority, and sanctity, but weakly or not at all related to the individualizing foundations. SDO (like economic orientation in the US, and right-orientation in Europe) is negatively associated with care and fairness, and more weakly than RWA associated with the binding foundations. The effect size differences are larger than those with the corresponding PSP items, ranging from Δr = .22 to .44.

Note on sample sizes

Using the R script provided by Gelman and Carlin (2014), to detect our chosen smallest relevant difference of r = .10 with a power of the recommended 80 %, one would need a sample size of 783. The median sample size in the samples included in this meta-analysis was 250, and only eleven samples—including the YourMorals dataset—had a sample size larger than required. Instead, with a n = 250 and thus SE = sqrt((1-0.1^2)/(250-2)) = 0.06318, we have probability of around 35 % to obtain a critical value for p to be .05 or less, and if a correlation would be found significant, it would likely be inflated by 67 % (Gelman & Carlin, 2014). The median effect size of all the main study-level analysis estimates was r = .288, for which the required sample size to reach 80 % power is 92, meaning that as much as three samples included in this meta-analysis would not have reliably detected half of the primary effects reported in this article. The median sample size of 250 would mean 80 % power for detecting effects larger than r = .177, while most of the estimates

for the care-PSP associations, and many differences in the moderator analyses, were below that. Our conclusion is that with the relatively small effect sizes found for MFQ-PSP associations, the sample sizes of studies should be increased considerably to reliably detect them for all MFQ factors. The minimum sample size should be from 194 (for an effect of $r = .2$) to 780 (for an effect of $r = .1$) to reach a power of 80 %.

Sensitivity analysis on different uses of MFQ

We ran study-level meta-analyses to find out whether different uses of the MFQ measure would have influence on the results: we tested (a) employing vs. omitting the attention check items ("Whether or not someone was good at math" and "It is better to do good than to do bad."), (b) full 30-item version vs. the brief 20-item version, and (c) using the whole MFQ scale (30 items) vs. use of relevance items only (15 items) vs. use of judgment items only (15 items). The details can be found in the Supplementary materials S2. The use of the attention check had, at highest, effects of $\Delta r = .060$, for loyalty, and $\Delta r = .073$, for sanctity, both decreasing the strength of the correlation ($k = 71$). The differences due to using the 20-item version instead of the full 30-item version were likewise small but also without a consistent direction, the largest difference again present in sanctity ($\Delta r = .068$; all other Δrs < .03). The practice to only use one half of the questionnaire had a slightly larger difference in correlations between binding foundations and a PSP measure. When using only the relevance items, the PSP-authority correlations were $\Delta r = .123$ smaller ($\Delta r = .075$ for sanctity and .087 for loyalty; for care and fairness Δrs < .042) than when the whole MFQ was in use. The consistent diminishing effect perhaps explains part of the differences between estimates in this meta-analysis and those found in Klein et al. (2018). Using only judgment items did not have a similar effect (for all MFQ factors, Δrs < .026).

Publication bias

Funnel plots regarding each MFQ-PSP measure pair form a generally symmetrical pattern around the average, with some effects outside the funnel on both sides, and the YourMorals dataset at the border (as discussed above). We present the funnel plot for the care-PSP effects as an example in Figure 12; the plots for the other MFQ factors, available in the Supplemental Figures S3F1 thru S3F4, do not notably deviate from it.

P-uniform* L-values range from 0.010 (sanctity-right) to 2.50 (loyalty-right), all *p*s > .25, suggesting that there is no selection based on the significance of the examined correlations. Results show no evidence for the presence of publication bias for any of the PSP measures or foundations (see full table in Supplemental Table S3T3).

For the Vevea & Woods (2005) selection method sensitivity analyses, we specified two sets of the *p*-value intervals and weights, assuming a moderate and severe publication bias (see details in Supplementary S2). The *p*-value intervals were specified to intervals from 0 to .05, from .05 to.10, and from .10 to 1. The weight assigned to the first interval was 1, meaning that it was assumed that all statistically significant results were published. For the severe adjustments, the weight for the second interval was 0.75, as proposed by Vevea and Woods (2005). For the last interval we assigned a weight of .30, meaning any correlation with an associated *p*-value larger than 0.10 was unlikely to be published. Even with these strong assumptions the difference between adjusted and unadjusted estimates was at maximum $\Delta r = .02$ (see details in Supplemental Table S3T4). To provide more breadth, we also tested the moderate adjustments (weights: probability of $p < .05 = 1$; probability of $.05 < p < .10 = .80$, and probability of $p > .10 = .50$), for which the differences were predictably smaller. We conclude that these analyses indicate no evidence for publication bias.

Discussion

We reviewed the literature on the association between moral foundations and political orientation, first presented in the influential article by Graham et al. (2009) entitled "Liberals and conservatives rely on different sets of moral foundations" and repeated in Graham et al. (2011). We elaborated this claim by identifying four assumptions that have been widespread in the reviewed research: assumptions of Direction, Magnitude, Unidimensionality, and Universality. Our meta-analysis covers 89 samples, 605 effect sizes, and 33,804 independent participants—in addition to 192,870 participants from Graham and colleagues' large internet-based YourMorals dataset. The data was limited to both moral foundations and political orientation measured with the most popular methods, the MFQ and individual self-placement items, that were the only ones providing enough studies to compare in a meta-analysis.

The results, with averaged political orientation measures and no distinctions made between samples or their origin, indicate that although we found moderate heterogeneity, the assumption of Direction is robust. With a few exceptions, the meta-analytic estimates suggest that care and fairness are generally negatively, and loyalty, authority, and sanctity, generally positively related to conservative political orientation. The relative Magnitudes between different moral foundations indicated that both the individualizing foundations care and fairness had weaker relationship with political orientation than the binding foundations—loyalty, authority, and especially sanctity. The association between care and political orientation was consistently the smallest in size, suggesting a broad agreement across the political divide on moral questions about harming and caring about others. However, the moderator analyses revealed several notable exceptions to the above results.

Political orientation associated with moral foundations in two dimensions

Contrary to the assumption of Unidimensionality, the evidence supports a bidimensional model on political orientation (e.g., Duckitt & Sibley, 2009; Feldman & Johnston, 2014), but also the conclusions that these dimensions are—sometimes strongly—correlated rather than independent (e.g., Jost et al., 2009). Although the differences were small, the consistency between different analyses drew a clear pattern. The associations described above appear to be more typical to the social, rather than economic political orientation. The latter was more equally (Δr range of .1, rather than .3), and less strongly (max $|r|$ = .25, rather than .45), associated with all moral foundations. This pattern was also supported by the alternative operationalization of political orientation as right-wing authoritarianism (RWA) corresponding to the social dimension, and social dominance orientation (SDO) corresponding to the economic dimension, as theorized in the dual-process model of ideology (Duckitt & Sibley, 2009). Although RWA and SDO were used as external validity criteria for authority and fairness in construction of the MFQ, respectively (Graham et al., 2011), our results point to more general associations across all five foundations, repeating the pattern of social and economic political orientation with much stronger associations and differences (cf. Altermatt et al., 2016; Dimdins et al., 2016; Federico et al., 2013; Hadarics & Kende, 2018). The difference between the dimensions is largest in association with sanctity, which has been found to be the most divisive moral foundation in earlier studies (e.g., Koleva et al., 2012), but that has a similar relationship to the economic dimension than other moral foundations. Both the generally weak relationship between care and political orientation and the different pattern of associations of moral foundations in relation to the economic orientation also support keeping the five moral foundations separate over the practice of collapsing them into two larger factors.

These findings are further moderated by political culture. Study-level analyses indicated that the associations of moral foundations to political orientation in the US show bidimensionality only when social and economic orientation were explicitly named, and were practically unidimensional when the self-placement items were labeled either as “left-right” or “liberal-conservative”. In contrast, European studies did not use explicitly named social and economic items, but the bidimensional pattern was apparent with left-right and liberal-conservative, where the former corresponded to the economic, and the latter to the social orientation. More differences were found from individual countries, most pronounced in the data from Latvia, where political orientation labeled as conservatism or right-orientation had associations of different sign (positive for conservatism, negative for right-orientation) to sanctity, and perhaps authority. These results are consequential, as they imply that in some cases even the Direction assumption may be violated (cf. Malka et al., 2019). Although the finding is based on very limited data, it is consistent with other work suggesting that conserving the traditional heritage may in postcommunist countries also mean protecting the communist social order that once represented the political left (Duriez, Van Hiel & Kossowska, 2005; Thorisdottir et al., 2007; Zarycki, 2000). The meaning of the labels will always vary somewhat due to historical and political considerations (Arian & Shamir, 1983; Inglehart & Klingemann, 1976; Piurko et al., 2011), and the meaning of the left-right dimension in postcommunist countries may differ significantly from the Western usage (Gunther & Kuan, 2007; Piurko et al., 2011; Whitefield, 2002), leading to generally weaker and even opposite-sign associations between left-right placement and psychological predispositions (e.g., Kossowska & Hiel, 2003; for a review, see Federico & Malka, 2018). It remains a highly intriguing question for future research if our results pertaining to Latvia can be replicated in other postcommunist nations.

The finding that the associations between morality and politics are dependent on the culture is at odds with the cultural sensitivity analysis in Many Labs 2 (Klein et al., 2018), which did not find differences between samples from WEIRD and non-WEIRD countries (although see our criticism in the Introduction questioning whether the comparison was really about WEIRD vs non-WEIRD). This means that in other countries, with political cultures farther from the US and Europe, the relationships may vary even more (cf. Pan & Xu, 2017, where the three political dimensions in China). This notion is supported by the fact that the data from New Zealand, another Anglosphere country, showed a pattern most resembling that from the US.

The dependence of bidimensionality on political culture may also be reflected against the classic distinction between "easy" and "hard" issues (Carmines & Stimson, 1980; Pollock, Lilie, & Vittes, 1993). According to earlier research, economic questions are considered to be "hard" questions, with often no immediate connection between a particular economic policy and particular beliefs or feelings that stem from some underlying characteristics of the individual, such as moral foundations. Rather, such associations are thought to be constructed by the political elite and communicated to the populace. In contrast to "hard" issues, "easy" issues are non-technical and ends oriented, and the associations between the suggested policies and the feelings that they elicit are intuitive. It has been shown that sanctity, which includes two items about disgust, is linked to the most visceral responses (Inbar et al., 2009; Russell & Giner-Sorolla, 2013), arguably making sanctity-related issues "easy"—in line with our findings that the strongest effects across samples were for sanctity. The pattern of results more generally supports the idea that the conservative-liberal labeling of political orientation may lead to more universal or generalizable associations with moral foundations, as policy preferences pertaining to social or moral issues; i.e., "easy" issues that elicit gut responses, may across cultures and historical contexts show more consistent associations

with moral foundations (cf. Hibbing et al., 2014). According to this view, how moral foundations are associated with economic preferences may depend more on cultural and historical factors, meaning that the correlations with political orientation as measured with the left-right label may vary much more across cultures (Bauer et al., 2017; Piurko et al., 2011). However, this would imply that in cultural circumstances where the social and economic dimensions are equated—e.g., the US, at the top of the list in recent comparison of 99 countries on how correlated the two dimensions are (see Malka et al., 2019)—the bidimensionality in relation to moral foundations should be the weakest. Yet, our results indicate that when explicit social and economic labels were used, the same bidimensional pattern emerged.

In addition to cultural differences, previous research shows that a major factor for whether political space appears as uni- or bidimensional is political engagement (Feldman, 2013; e.g., Kinder & Kalmoe, 2017). If the associations by the economic policies and the beliefs or feelings stemming from psychological dispositions have to be constructed by the political elite, they are more likely learned by those who follow the elite political discourse closely—are more engaged with politics. However, our results are at odds with this, as political interest showed no effect on bidimensionality. It is unclear whether this is due to the difference between political engagement and political interest (despite it being used as a factor within political engagement measures, one can be interested without being engaged), the lack of data (the analyses only included three studies), some difference between bidimensionality and how it links to moral foundations or political interest —or whether there is some "easy" component behind economic orientation as well.

The influence of sampling and political interest

The previous large-scale replication effort (Klein et al., 2018) indicated that in general, when real effects are found, they are quite robust to minor changes in experimental procedures or

sampling details. Our results mostly support these findings: the results appear rather robust to differences in analysis methodology, the differences between MFQ long and short form, and the use of attention check items (some differences were found between using the whole MFQ and only relevance or judgment items, however). Although there were some differences between the representative to MTurk to convenience sampling methods, they were generally small. The YourMorals sample remains a notable exception—its estimates yielding almost double to more than double the effect sizes compared to other sampling methods. Although the sample size of the YourMorals dataset is an order of magnitude higher than all the independent studies together (including datasets from large-scale projects such as NZAVS and ANES), the consistency across the independent studies makes the estimates from the YourMorals dataset suspect.

A possible explanation is the amount of self-selection, which may result in a skewed sample people particularly interested in the topic of the research espousing more extreme attitudes or beliefs than people who are indifferent about politics and moral theories. Self-selection could be expected to be smallest when the items on morality and political issues are run along with a large number of items on other topics, such as in representative samples and large-scale multi-lab replications (Klein et al., 2018, with clearly smaller effects). Self-selection would be expected to be somewhat higher typical student-based convenience samples, and larger still in MTurk samples where the respondents choose the questionnaires they respond to—notably, MTurk effects were typically the closest to YourMorals data effects, and may involve some inflation as well. In contrast, the YourMorals sample respondents have themselves actively sought the questionnaires, indicating strong self-selection. We investigated this issue with a series of individual-level analyses using political interest as a moderator, and found that the associations of political orientation to care and fairness present in the YourMorals sample were dependent on political interest, while the associations to binding

foundations were practically independent of it. In addition, this pattern appeared to hold to a smaller extent for the representative and MTurk samples as well. Returning to the distinction between the "easy" and "hard" political questions, it is notable that political interest moderates the associations of the PSP to care and fairness factors, with items more related to the economic dimension.

We also investigated whether the demographical skews influenced the differences between YourMorals and other samples. We found that, despite the common link between education and liberalism, differences in education did not explain the sample type differences. Neither did age (cf. Cornelis, Van Hiel, Roets, & Kossowska, 2009) or gender, although we did find that the negative association between conservatism and individualizing foundations was somewhat smaller for women, indicating that both liberal and conservative women are more care-oriented and the relative unimportance of care is related more specifically to conservative men (cf. Koleva et al., 2014). Religion—compared between Christian, No religion, and Other religion groups—did not explain the sample type differences, and while Christians endorsed binding foundations at higher average than the no or other religion groups, there was little interaction effect with conservatism. Likewise, race failed to explain the sample type differences. However, race showed a small tendency diminishing the sample type differences for Black and Hispanic respondents, and together with political interest (that had a stronger effect than race) the individualizing-political orientation association dropped to zero and the binding-political orientation associations decreased notably, mostly erasing the effect inflation in the YourMorals sample. This suggests that the higher estimates in the YourMorals sample are due to a combination of self-selection and biased racial demographics.

Other findings

In the representative samples, arguably giving us the least biased estimates for the general population, and its subset of Black respondents, all associations between moral foundations and

political orientation were close to zero ($r <= .1$), and further reduced by higher political interest. These results dovetail those of Davis et al. (2016), who showed that the associations between the binding foundations sanctity and authority with conservatism were weaker among Black people than White people. Importantly, the Black people in Davis's as well as the independent samples here were almost exclusively part of US samples. During the last decades a more ideological and issue-based form of partisanship has emerged in US politics, meaning that there is a strong relationship between partisanship and liberal-conservative self-placement (Bafumi & Shapiro, 2009). However, for Black people, ideology is a much weaker organizing structure for political attitudes than for White people (Chen, 2017; Philpot, 2017). While Black people may identify as conservatives, this is less indicative of their political—and likely, moral—stances, leading to weaker relationship between our measures of political orientation and morality. As these phenomena are molded by the historical factors operating specifically in the US, we have no reasons to believe that they would directly generalize to very different cultural environments.

Numerous studies have used only half of the MFQ, the relevance or judgment part, instead of the whole instrument (e.g., Klein et al., 2018). It has been reported several times that the internal consistencies are naturally lower when the number of items decreases, but also that the judgment subscale is worse in this regard than the relevance subscale (e.g., Bobbio et al., 2011; Zhang & Li, 2015). However, we found that for the judgment subscale did not perform differently to a relevant extent as compared to the full MFQ, while when measured by relevance subscale only, the loyalty, authority, and sanctity foundations showed differences ranging from $\Delta r = .075$ to .123. It could be that the nature of the task—evaluative judgments vs. self-reflection—is crucial (cf. Norris, Larsen, Crawford, & Cacioppo, 2011, for differences in emotional evaluation in similar tasks). Another explanation might be that the most important items happen to be in the judgment subscales. In the

domain of personality psychology, it has been suggested that single items, in the prediction of relevant outcome measures, often outperform the broader (underlying) traits that these items are ostensibly indicators of (Mõttus, 2016; Seeboth & Mõttus, 2018). It could be that only a few items of the MFQ are responsible for the associations to political orientation, and that these items would outperform the five foundations in predicting political orientation. Moreover, it could be the same items or different items that in different data sets underlie the associations. Sorting this out requires further investigations.

Risk of bias

A priori, we expected little bias due to publication pressures, because the meta-analyzed correlations were not the focus in any of the included articles and are therefore unlikely to influence the decision to publish or not. The funnel plots did not suggest a skew, and the *p**-uniform analyses showed that as a collection, the meta-analyzed effect sizes did not portray suspicious distributions accumulating around the common cutoff of significance, $p = .05$. Sensitivity analysis that assumed a severe bias and adjusted a set of results to correct them likewise produced minuscule adjustments, again indicating no publication bias.

The quality of studies can be influenced by non-attentive or uncooperating respondents, the effect of which is typically combated with attention checks. We tested the differences in estimates from studies that used vs. did not use the attention check items included in standard MFQ. The differences were small (max $\Delta r = .076$) and were concluded not to bias the results to a relevant extent.

Strengths and limitations

This review provides a qualitative and quantitative overview of the current literature on the association between moral foundations and political orientation. The review followed a formal

protocol and it screened over a thousand records, ultimately including 67 publications and 89 samples, with over 30,000 independent respondents in addition to the 192,000 respondents from YourMorals, from thirteen different countries. Because we limited the review to very specific measures and ruled out samples with explicitly or potentially influencing experimental manipulations, the quality of data is very homogenous. The use of multi-level models, robust variance estimation, and sensitivity analyses reduces potential problems introduced by dependent data and influential outliers. In addition, modeling on level of both studies and individual participants gave us flexibility to pursue research questions that would have otherwise been impossible to investigate.

A quantitative analysis is only as good as its measures, and the homogeneity of the measures is not only a strength, but also a limiting factor. Although focusing on these measures is justified due to how common they are, almost to the extent of exclusivity, both the MFQ and PSP measures have been criticized on theoretical and methodological grounds. For political orientation, the single-item self-placement scale with seven steps, from “very liberal” to “very conservative”, is the most typical measure used, both generally in political psychology (Jost, 2006) and in our data. However, there is some disagreement on what it measures. Research on symbolic and operational ideology suggest a disparity between how people identify politically and what are their actual attitudes concerning practical policy issues (Ellis & Stimson, 2009). The former may guide how they respond to a self-identification measure, while the latter may be more relevant to questions or tasks regarding morality. This may also be one reason for the findings about the differences between the racial groups (Philpot, 2017). A self-placement item also relies on the understanding of the respondent to map the variety of their political leanings on the scale, and that understanding clearly depends on the political culture in which they have learned those terms (e.g., Bauer et al., 2017). We

looked at two qualifications on the use of the common political orientation measures (different descriptors for the dimension and different geographical regions representing different political cultures), but it is unlikely that these would be the only ones that matter. However, these concerns are somewhat mitigated by the fact that our analyses utilizing correlations calculated to RWA and SDO multi-item measures instead of single-item self-placement scales at least support our conclusions about bidimensionality.

Similarly, Moral Foundations Questionnaire has been criticized on, for example, poor indices of internal consistency, especially for translated items, and the studies examining the structural validity of MFQ have noted model fits poorer than typically recommended. The problems of low validity and reliability may lead to large error rates regardless of large sample sizes (Westfall & Yarkoni, 2016). Low alphas indicate that the items are not very well correlated, i.e. that a substantial portion of the variance is item-specific instead of attributable to a general factor (Cortina, 1993). However, it should also be noted that due to the small number of items (six, or four in the short form) per subscale, alpha underestimates reliability (Tavakol & Dennick, 2011). To an extent, this may mean that the factors are not measuring a single construct but rather several interlinked processes, although arguably, this can be also desirable, depending on the aims of the measure ("conceptual breadth instead of internal consistency", as stated in Graham et al., 2011). However, if the alphas would be considerably smaller in the translated studies, this would likely indicate increase in item-specific random error rather than conceptual breadth. It is possible that some portion of specifically our findings about cultural differences may be affected by this, as demonstrated by the alpha differences between samples. Cross-cultural studies have previously criticized that while the MFT has been specifically created to cover morality in a more broad manner than WEIRD views alone, MFQ still has many items that seem to rely on the Western- or

US-centered cultural context. Examples include references to God that fit poorly to non-Abrahamic religions, and strongly linking loyalty to patriotism, which may be problematic in countries with less cultural importance on country and more, e.g., on family and personal relationships. The research on cultural differences is sparse, and it is not clear to which extent the current results can be expected to generalize to most of the humanity outside the thirteen origin countries of the studies—overrepresenting the US and Europe—or to less represented groups within these countries. Clearly, more detailed investigations on the cultural psychology and psychometrics of the MFQ are warranted. However, the psychometric details do not change the theoretical and empirical basis behind MFT. Even as it may be questioned, e.g., just like political orientation, how the respondents understand particular questions, they still link to the scholarly traditions of Kohlberg, Gilligan, Shweder, and others. Alternative models of morality, regardless of the exact formulation (such as Curry et al., 2019; Janoff-Bulman & Carnes, 2013; see below), necessarily tap into the same moral issues and probe the same political cultures; thus, there is little reason to believe that other constructs and other measures would give substantially different results regarding the main findings.

It should be also noted that MFQ is not the only measure for moral foundations. Although alternatives were not included here due to their rare use, there are others already available, such as Moral Foundations Sacredness Scale and the Moral Foundations Dictionary that were already used by Graham and colleagues (2009), establishing the associations to political orientation. Others have constructed new instruments (e.g., Clifford et al., 2015) or ways to research moral foundations (e.g., Feinberg & Willer, 2015; Frimer et al., 2013). In addition, Graham and colleagues are constructing a new version of the MFQ, hopefully covering these issues that have been long known.

The vast majority of the included studies used sampling that limits our ability to generalize beyond students and MTurk volunteers. While we used representative samples as a benchmark for

more probably true estimates of the effect sizes, we could only find four such samples, originating from different countries and likely collected using different methodologies. Furthermore, it may be relevant to question the representativeness of a sample collected from internet panels, as the articles do not always report sampling methods beyond "nationally representative" (Kivikangas et al., 2017; Nilsson et al., 2016). For example, Chang and Krosnick (2009) suggest that recruitment from the panel by invitation that describes the content of the survey may lead to selection by topic interest, essentially the same issue we partly responsible for the inflation of correlations in the YourMorals sample. In addition, the original recruitment of the volunteers to the panels may reflect the common biases in sampling and selecting for more open and agreeable people than the general population. Nevertheless, the widely used databases ANES and NZAVS can be considered some of the highest quality data sources available, and that they did not stand out as outliers is reassuring.

Practical and theoretical implications

First, we emphasize that the importance of our results is not limited to a particular theory of morality (MFT) or a particular measure (MFQ). The take-home message is more general and applies to all work in which individual differences in how we think and feel about moral and ethical issues is associated with how we think and feel about political issues. That is, these associations will vary not only culturally, showcased e.g. by the US-Europe differences and the special case of Latvia, but also subculturally, showcased e.g. by the US Black respondents, as well as being dependent on how the measures are understood in the local culture. Moreover, individual differences, such as political interest, will influence these associations. Yet, our findings about the stability and robustness of the associations suggest that, when certain facts about the political culture and context of use are recognized and taken into account, the links between certain types of morality and politics appear to be rather predictable. This can be considered surprising in light of the on-going debate on how

exactly conservatism should be understood, where one argument has been that conservatism is an improvisation for defending hierarchy and privilege—i.e. a reaction that changes across time and space in response to the movement it opposes (Lilla, 2016; Robin, 2011)—rather than an ideology that opposes revolutionary social change based on a certain view of human nature. Our results suggest that both dimensions, social and economic, do have a somewhat uniform moral meaning, in that they appear to show moderately strong and robust correlations with the moral foundations. This may be interpreted as supporting the view in which political orientation is rooted in individual differences in more basic psychological traits, that predispose some individuals to be more willing to accept particular views on the social environment (traits such as threat sensitivity, complexity and uncertainty intolerance, and openness to new experiences; Jost et al., 2003; Jost & Amodio, 2012; see also Duckitt & Sibley, 2009). In addition, this challenges the implicit view that it is specifically conservatism that requires explanation, as very few studies aim to explain liberalism (cf. Duarte, Crawford, Stern, Haidt, Jussim, & Tetlock, 2015).

On an even more general level, our results contribute to the discussion on the bottom-up vs. top-down processes through which individuals come to adopt political preferences. In the bottom-up view, the individual differences in psychological predispositions such as moral foundations are thought to manifest as tendencies to adopt certain political preferences (Federico & Malka, 2018; Feldman & Johnston, 2014; Jost et al., 2009). A top-down perspective would, on the other hand, emphasize the role of communication by political elites (e.g., political leadership, party politics, political scientists, mass media). However, not everyone will receive the messages of political discourse to the same degree, and only those who are the most politically engaged will adopt a coherent set of attitudes (e.g., Converse, 1964; Zaller, 1992). Our results follow this line of thought by suggesting that the politically interested will adopt political preferences more consistent with

their moral preferences. The cross-cultural and subcultural variability that we observed in turn emphasizes the role of top-down processes, in which political discourse come to play in shaping the associations between moral and political preferences.

Going to the practical implications, the finding that the Direction of the associations between moral foundations and political orientation as well as the relative Magnitude between the foundations is quite robust means that the claim "liberals and conservatives rely on different moral foundations" holds for more general purposes. Unless the focus is specifically in the economic dimension, care is least associated with political orientation, indicating the greatest agreement across the political divide, while the opposite is true for sanctity. Furthermore, when violations of assumptions of Direction, Magnitude, Unidimensionality, or Universality were found, they were generally small to moderate in effect size. However, as is known, the effects generally examined in psychology are not large, so specific details may be more consequential than their absolute effect sizes lead to expect. Research focusing on a particular moral foundation, specifically economic political orientation or topics related to it, or a population markedly different from the most common one (i.e. White male, from the US or Western Europe, interested in politics; see below) should be aware of the differences revealed in this meta-analysis.

The differences between sample types have implications for comparisons to the YourMorals data and sampling in general. One side of this is that the articles that directly use YourMorals data or do comparisons between it and other data are directly affected by the findings that the correlations based on YourMorals data seem to be heavily inflated. For instance, Kim and others (2012) compared their rather weak correlations (ranging from r = -.08 to .25, for care and authority, respectively) from South Korea to the inflated YourMorals correlations representing the US, leading to conclusions about larger differences than would be accurate, according to our independent

studies. The findings about both sample types and the effect of context in which the sample was collected (e.g., race, political culture) give new weight to the warning by Davis and colleagues (2016) that the associations between moral foundations and political orientation do not necessarily generalize to samples very different from the original YourMorals data. Although we did not find notable effects for other demographics and individual differences than race and political interest, the case of Latvia alone indicates that there may be other important factors. Unexpected patterns may be found in parts of East Europe or elsewhere in the world (Malka et al., 2019; Thorisdottir et al., 2007), or non-mainstream political groups such as nationalist and anti-establishment egalitarian populists (Koster, Achterberg, & Waal, 2013), or libertarians, who endorse the moral foundations in unique patterns (Iyer et al., 2012).

The generally modest size of findings echoes calls for caution by many other studies warning that the current empirical research shows inadequate power in relation to the effect sizes it attempts to capture (Colquhoun, 2014; Gelman & Carlin, 2014; Klein et al., 2018; Open Science Collaboration, 2015). Depending on the expected effect size of interest, future studies investigating moral foundations and political issues should collect a sample size of, at minimum, $n = 194$ to detect most of the effects, to as much as 780 to ensure finding the smallest effects as well. In addition, we note that for investigations of particularly small effects, the attention check provided within MFQ and using the whole 30-item instrument should increase the quality of responses.

In addition to the sampling issues, the differences revealed in our analyses also have practical implications for designing new research, such as when the study design is (partially) dependent on the assumed relationship between moral foundations and political orientation. For example, Frimer and colleagues (2013) investigated whether experts and common people with different political orientation rely on the same moral foundations when judging the moral character

of public figures, and found that both liberals and conservatives mainly judge morality based on care, fairness, and sanctity. Considering that care and fairness (although not sanctity) are stronger points of agreement on the social dimension, it may be that more disagreement on morality of at least some public figures could be found if the raters were selected based on the economic dimension instead, especially outside the US. Similarly, our results indicating that political interest is a predictor of associations between political orientation and individualizing foundations give raise to concerns about sampling methods that capitalize on political interest of volunteers. While interest in morality and its relationship to politics can be of great help in recruitment, especially before elections, it may skew the sample and lead to erroneous conclusions. If such methods are employed, it is recommended that variables such as political interest are used to at least assess the severity of the bias.

Our results also suggest that the practice of collapsing the five foundations into the two higher-order factors would be more suitable when examining topics related to social political orientation, or issues without distinguishing between the social and economic dimensions. As sanctity is less and care and fairness more related to the economic dimension, it suggests that in topics related more to the economic issues the use of two factors could be misleading. In the same vein, not separating care and fairness may bias the association the individualizing higher-order factor would have with political orientation, as care appears to be more variable and even have correlations opposite to fairness with political orientation in certain samples.

Future directions

The current study, our interpretations on its results, and its limitations leave open a number of questions that we recommend for investigation in future research. As MFT and its primary instrument, MFQ, are used across the globe, an important question is whether the results hold in

different countries and political cultures. After all, the countries we studied are still mostly very similar to the US in terms of the relevant political issues, while some completely other issues (such as personal freedoms and nationalism in China; Pan & Xu, 2017) may be more relevant in different political cultures. Moreover, the issues the MFQ may have with representing morality in culturally diverse samples call for more cross-cultural studies, with data more representative of local contexts and without the biasing effect of self-selection. The possible divergence between self-identification, especially in polarized political cultures, and attitudes on practical issues, as well as the influence of political engagement beyond political interest, also suggest new fruitful avenues for future research.

As a relatively new theory, it is not clear how well MFT represents the underlying psychological structure of morality, and several independent scholars have criticized the idea of moral plurality or its taxonomy, offering alternative models. Some of these alternatives present aspects of morality that might have a considerable influence on the morality-politics associations studied here. Gray and colleagues (Gray & Keeney, 2015; Gray, Young, & Waytz, 2012; Schein & Gray, 2015, 2018) have argued that the five separate foundations are artefacts resulting from faulty methods and that all moral judgments are really perceptions of a harm (inverse of care) between the moral dyad, agent and patient. It is unclear how focusing on the different forms of harm (e.g., loyalty as a harm to own group) rather than separate foundations would change the results or the conclusions, but the concept of the moral dyad suggests an interesting lens to issues where, in addition to what is being done, it often politically matters who is doing it and to whom—a distinction mostly missing from the MFT. How would the relative agreement about care and harm change, if when measuring moral differences, the agents and patients related to moral issues were taken into account? Some studies have suggested that the moral expansiveness is narrower for conservatives, who consider morality more in terms of close ones and ingroups, than for liberals,

who are more universalist in their morality (Crimston et al., 2016; see also Schwartz, 2007). Would liberals be more willing to espouse the binding foundations if they were more about self-sacrifice on behalf of family (loyalty) and cherishing beautiful traditions (authority) and purity of nature (sanctity)? At least one study has suggested obedience to authority is not specific to conservatives, as liberals are as likely to obey if the authorities are such that liberals consider them more relevant (e.g., civil rights leaders or environmentalists; Frimer, Gaucher, & Schaefer, 2014).

In the theoretical alternative put forth by Janoff-Bulman and Carnes, the Model of Moral Motives, the difference between liberals and conservatives is not in their endorsement of binding morality but in whether the group-based morality focuses on proscriptive (prevent the bad) or prescriptive (promote the good) motives (Janoff-Bulman & Carnes, 2014, 2016)—another distinction that matters a great deal in politics, but that MFT currently overlooks.

Recently, a new theory challenged MFT arguing that the function of morality is to promote cooperation, that game theory can help identify distinct types of cooperation, and that these types can be empirically found to be considered universally moral (Curry et al., 2019; Curry, Mullins, & Whitehouse, 2019). Curry and colleagues propose seven types of cooperation, partially overlapping with moral foundations, in their Morality-as-Cooperation theory. They argue that the instrument they developed, the MAC-Q, is psychometrically better than the MFQ, and that comparison between the two supports seven types of cooperation over five moral foundations (Curry, Chesters, et al., 2019). Despite the different roots of the theory, the seven types of moral cooperation offer an alternative set of moral domains resembling foundations (at least some of them already being considered as potential foundations by Graham and others, 2013). Curry and others mention that "because the value of these different types of cooperation can vary independently, moral values in each domain will vary independently too" (Curry, Mullins, & Whitehouse, 2019, p.63), but they do

not report how these values are associated with political orientation. It would be interesting to see if helping kin and one's group, reciprocating, and deferring to superiors have similar relationships to dimensions of political orientation as loyalty, fairness, and authority, and to what extent the types of cooperation neglected by moral foundations are politically relevant.

Conclusions

In our meta-analysis we found that the claim that liberals and conservatives rely on different sets of moral foundations in general holds, albeit with notable exceptions. That is, care and fairness are negatively, and authority, loyalty, and sanctity are positively correlated with political conservatism. Care and fairness had generally weaker associations to political orientation than the three binding foundations. The association between care and political orientation was the smallest, suggesting that caring and avoiding harm is more widely regarded as important across the political divide. However, heterogeneity was moderate, and the assumption that moral foundations have unidimensional associations with political orientation was not supported by the analysis. Although the differences were rather modest, the foundations were more strongly correlated with social than economic political orientation, sanctity in particular. Hence, the results indicate that researchers of morality should pay more attention on the political measures they use, and that collapsing the foundations to the two higher order factors is not warranted if the differences between social and economic political orientation are relevant. Universality assumption was also questioned, as differences were found across cultures, demographics, and sampling methods. Political labels had somewhat different meanings in the US and Europe, and associations between moral foundations and political orientation varied across countries and political cultures. We found that the widely used, large-scale YourMorals sample has considerably inflated effect sizes. This difference was not explained by most demographics (education, gender, age) but is likely related to self-selection and

confounding political interest, which may also influence popular MTurk sampling. However, moral foundations-political orientation associations were diminished considerably when examining Black, and to smaller extent, Hispanic population instead of White in US, or when examining samples from countries with different political history and culture. These differences have implications for designing studies and their sampling.

Apart from political orientation labels, methodological differences, such as long or short form of the questionnaire, caused very small differences at best. The largest differences were for the use of only the judgment vs. relevance items of the MFQ, raising concern about the validity of studies using only half of the questionnaire. The effect sizes generally were small to moderate, suggesting that in most studies the sample size is not adequately big to capture the effect reliably, and that future studies examining the relation of moral foundations and political orientation should have sample sizes of at least 200, or ideally, near 800.

In conclusion, the results of this meta-analysis both confirm the general relations of moral foundations and political orientation, but also raise some notable considerations for future studies regarding the methodology and theory.

References

References with asterisks were included in the meta-analysis. Note that articles using data from one of the included databases (YourMorals, ANES, NZAVS) were not used as primary studies separately.

Aloe, A. M. (2015). Inaccuracy of regression results in replacing bivariate correlations. *Research Synthesis Methods*, *6*(1), 21–27. https://doi.org/10.1002/jrsm.1126

Altemeyer, B. (2006). *The Authoritarians*. Retrieved from http://theauthoritarians.org/Downloads/TheAuthoritarians.pdf

*Altermatt, B., Martell, J., McCleary, L., Reinhart, M., & Nevils, B. (2015). *The Fairness Foundation is About More than Equality, Equity, or Care* [Poster]. Annual meeting of the Midwestern Psychological Association, Chicago, IL. Retrieved from https://vault.hanover.edu/~altermattw/research/MoralPsych/index.html

*Altermatt, B., Martell, J., Reinhart, M., Nevils, B., Brown, M., & Stine, E. (2016, January). *Applying a two-dimensional theory of political ideology to moral foundations theory*. Poster presented at the Annual Convention of Society for Personality and Social Psychology, San Diego, CA. Retrieved from osf.io/qc4vn

*American National Election Studies. (2010a, September 3). *2008-2009 Panel Study (Wave 6)*. Retrieved from https://electionstudies.org/data-center/

*American National Election Studies. (2010b, October 7). *Auxiliary File: 2008-2009 Panel Study (Wave 7)*. Retrieved from https://electionstudies.org/data-center/

Arian, A., & Shamir, M. (1983). The primarily political functions of the left-right continuum. Comparative Politics, 15(2), 139-158.

Aspelund, A., Lindeman, M., & Verkasalo, M. (2013). Political Conservatism and Left–Right Orientation in 28 Eastern and Western European Countries. *Political Psychology*, *34*(3), 409–417. https://doi.org/10.1111/pops.12000

Bafumi, J., & Shapiro, R. Y. (2009). A New Partisan Voter. *The Journal of Politics*, *71*(1), 1–24. https://doi.org/10.1017/S0022381608090014

*Baril, G. L., & Wright, J. C. (2012). Different types of moral cognition: Moral stages versus moral foundations. *Personality and Individual Differences*, *53*(4), 468–473. https://doi.org/10.1016/j.paid.2012.04.018

*Barnett, M. D., Öz, H. C. M., & Marsden, A. D. (2017). Economic and Social Political Ideology and Homophobia: The Mediating Role of Binding and Individualizing Moral Foundations. *Archives of Sexual Behavior*, 1–12. https://doi.org/10.1007/s10508-017-0989-2

*Barnett, M. D., & Hilz, E. N. (2017). The Psychology of the Politics of Rape: Political Ideology, Moral Foundations, and Attitudes Toward Rape. *Violence Against Women*, 1077801217708887. https://doi.org/10.1177/1077801217708887

Bassett, J. F., Van Tongeren, D. R., Green, J. D., Sonntag, M. E., & Kilpatrick, H. (2015). The interactive effects of mortality salience and political orientation on moral judgments. *British Journal of Social Psychology*, *54*(2), 306–323. https://doi.org/10.1111/bjso.12086

Bauer, P. C., Barberá, P., Ackermann, K., & Venetz, A. (2017). Is the Left-Right Scale a Valid Measure of Ideology? *Political Behavior*, *39*(3), 553–583. https://doi.org/10.1007/s11109-016-9368-2

*Black, J. E. (2017). *Unpublished dataset*.

*Black, J. E., & Barnes, J. L. (2017). Measuring the unimaginable: Imaginative resistance to fiction and related constructs. *Personality and Individual Differences*, *111*, 71–79. https://doi.org/10.1016/j.paid.2017.01.055

*Black, J. E., & Reynolds, W. M. (2016). Development, reliability, and validity of the Moral Identity Questionnaire. *Personality and Individual Differences*, *97*, 120–129. https://doi.org/10.1016/j.paid.2016.03.041

Bobbio, A., Nencini, A., & Sarrica, M. (2011). Il Moral Foundation Questionnaire: Analisi della struttura fattoriale della versione italiana [The Moral Foundation Questionnaire. Factorial structure of the Italian version]. *Giornale Di Psicologia*, *5*. Retrieved from http://www.psicotecnica.net/gdp/GiornaleDiPsicologia.2010.4.3.pdf

Borenstein, M., Higgins, J. P. T., Hedges, L. V., & Rothstein, H. R. (2017). Basics of meta-analysis: I^2 is not an absolute measure of heterogeneity. *Research Synthesis Methods*, *8*(1), 5–18. https://doi.org/10.1002/jrsm.1230

Bulbulia, J., Osborne, D., & Sibley, C. G. (2013). Moral foundations predict religious orientations in New Zealand. *PloS One*, *8*(12), e80224.

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., … Pfeiffer, T. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*, *2*(9), 637.

Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American Voter*. New York, NY: John Wiley.

*Cardinale, E. M., & Marsh, A. A. (2015). Impact of Psychopathy on Moral Judgments about Causing Fear and Physical Harm. *PLOS ONE*, *10*(5), e0125708. https://doi.org/10.1371/journal.pone.0125708

Carmines, E. G., & Stimson, J. A. (1980). The Two Faces of Issue Voting. *American Political Science Review*, *74*(1), 78–91. https://doi.org/10.2307/1955648

Carney, D. R., Jost, J. T., Gosling, S. D., & Potter, J. (2008). The Secret Lives of Liberals and Conservatives: Personality Profiles, Interaction Styles, and the Things They Leave Behind. *Political Psychology*, *29*(6), 807–840. https://doi.org/10.1111/j.1467-9221.2008.00668.x

Chang, L., & Krosnick, J. A. (2009). National surveys via RDD telephone interviewing versus the internet comparing sample representativeness and response quality. *Public Opinion Quarterly*, *73*(4), 641–678. https://doi.org/10.1093/poq/nfp075

Chen, E. E. (2017). *The Importance of Ideology varies across Sociocultural Contexts* (PhD Thesis). UC Irvine.

Cheung, M. W.-L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods*, *19*(2), 211–229. https://doi.org/10.1037/a0032968

Clifford, S. (2017). Individual Differences in Group Loyalty Predict Partisan Strength. *Political Behavior*, *39*(3), 531–552. https://doi.org/10.1007/s11109-016-9367-3

*Clifford, S., Iyengar, V., Cabeza, R., & Sinnott-Armstrong, W. (2015). Moral foundations vignettes: A standardized stimulus database of scenarios based on moral foundations theory. *Behavior Research Methods*, *47*(4), 1178–1198. https://doi.org/10.3758/s13428-014-0551-2

Coburn, K. M., & Vevea, J. L. (2016). *Weightr: Estimating weight-function models for publication bias in r*. R package version 1.0.0. Retrieved from https://CRAN.R-project.org/package=weightr

Cochrane, C. (2010). *Left/right asymmetries in a multidimensional universe: Citizens, activists, and parties* (University of Toronto). Retrieved from https://tspace.library.utoronto.ca/handle/1807/24414

*Cohen, T. R., Panter, A. T., Turan, N., Morse, L., & Kim, Y. (2014). Moral character in the workplace. *Journal of Personality and Social Psychology*, *107*(5), 943–963. http://dx.doi.org/10.1037/a0037245

Colquhoun, D. (2014). An investigation of the false discovery rate and the misinterpretation of p-values. *Royal Society Open Science*, *1*(3), 140216. https://doi.org/10.1098/rsos.140216

Converse, P. E. (1964). The nature of belief systems in mass publics. In D. E. Apter (Ed.), *Ideology and discontent* (pp. 206–261). New York, NY: Free Press.

Converse, P. E. (2000). Assessing the capacity of mass electorates. *Annual Review of Political Science, 3*(1), 331–353.

Cornelis, I., Van Hiel, A., Roets, A., & Kossowska, M. (2009). Age Differences in Conservatism: Evidence on the Mediating Effects of Personality and Cognitive Style. *Journal of Personality*, *77*(1), 51–88. https://doi.org/10.1111/j.1467-6494.2008.00538.x

*Cornwell, J. F. M., & Higgins, E. T. (2013). Morality and Its Relation to Political Ideology The Role of Promotion and Prevention Concerns. *Personality and Social Psychology Bulletin*, *39*(9), 1164–1172. https://doi.org/10.1177/0146167213489036

*Cornwell, J. F. M., & Higgins, E. T. (2014). Locomotion concerns with moral usefulness: When liberals endorse conservative binding moral foundations. *Journal of Experimental Social Psychology*, *50*, 109–117. https://doi.org/10.1016/j.jesp.2013.09.007

Cortina, J. M. (1993). What is coefficient alpha. *Journal of Applied Psychology*, *78*(1), 98–104.

*Crimston, D., Bain, P. G., Hornsey, M. J., & Bastian, B. (2016). Moral expansiveness: Examining variability in the extension of the moral world. *Journal of Personality and Social Psychology*, *111*(4), 636–653. https://doi.org/10.1037/pspp0000086

*Crone, D. L., & Laham, S. M. (2015). Multiple moral foundations predict responses to sacrificial dilemmas. *Personality and Individual Differences*, *85*, 60–65. https://doi.org/10.1016/j.paid.2015.04.041

Crowson, H. M. (2009). Are All Conservatives Alike? A Study of the Psychological Correlates of Cultural and Economic Conservatism. *The Journal of Psychology*, *143*(5), 449–463. https://doi.org/10.3200/JRL.143.5.449-463

Curry, O. S., Chesters, M. J., & Van Lissa, C. J. (2019). Mapping morality with a compass: Testing the theory of 'morality-as-cooperation' with a new questionnaire. *Journal of Research in Personality*, *78*, 106–124. https://doi.org/10.1016/j.jrp.2018.10.008

Curry, O. S., Mullins, D. A., & Whitehouse, H. (2019). Is It Good to Cooperate?: Testing the Theory of Morality-as-Cooperation in 60 Societies. *Current Anthropology*, *60*(1), 47–69. https://doi.org/10.1086/701478

Davies, C. L., Sibley, C. G., & Liu, J. H. (2014). Confirmatory Factor Analysis of the Moral Foundations Questionnaire: Independent Scale Validation in a New Zealand Sample. *Social Psychology*, *45*(6), 431–436. https://doi.org/10.1027/1864-9335/a000201

Davis, D. E., Rice, K., Van Tongeren, D. R., Hook, J. N., DeBlaere, C., Worthington, E. L., & Choe, E. (2016). The moral foundations hypothesis does not replicate well in Black samples. *Journal of Personality and Social Psychology*, *110*(4), e23–e30. https://doi.org/10.1037/pspp0000056

*Dawson, S. L., & Tyson, G. A. (2012). Will Morality or Political Ideology Determine Attitudes to Climate Change? *Australian Community Psychologist*, *24*(1), 8–25.

Day, M. V., Fiske, S. T., Downing, E. L., & Trail, T. E. (2014). Shifting Liberal and Conservative Attitudes Using Moral Foundations Theory. *Personality and Social Psychology Bulletin*, *40*(12), 1559–1573. https://doi.org/10.1177/0146167214551152

*Dimdins, G., Sandgren, M., & Montgomery, H. (2016). Psychological variables underlying political orientations in an old and a new democracy: A comparative study between Sweden and Latvia. *Scandinavian Journal of Psychology*, *57*(5), 437–445. https://doi.org/10.1111/sjop.12314

Dogruel, L., Joeckel, S., & Bowman, N. D. (2013). Elderly people and morality in virtual worlds: A cross-cultural analysis of elderly people's morality in interactive media. *New Media & Society*, *15*(2), 276–293. https://doi.org/10.1177/1461444812451571

Dollinger, S. J., & Leong, F. T. L. (1993). Volunteer Bias and the Five-Factor Model. *The Journal of Psychology*, *127*(1), 29–36. https://doi.org/10.1080/00223980.1993.9915540

Duarte, J. L., Crawford, J. T., Stern, C., Haidt, J., Jussim, L., & Tetlock, P. E. (2015). Political diversity will improve social psychological science. *Behavioral and Brain Sciences*, *38*. https://doi.org/10.1017/S0140525X14000430

Duckitt, J., & Sibley, C. G. (2009). A Dual-Process Motivational Model of Ideology, Politics, and Prejudice. *Psychological Inquiry*, *20*(2–3), 98–109. https://doi.org/10.1080/10478400903028540

*Dungan, J. A., Chakroff, A., & Young, L. (2017). The relevance of moral norms in distinct relational contexts: Purity versus harm norms regulate self-directed actions. *PLOS ONE*, *12*(3), e0173405. https://doi.org/10.1371/journal.pone.0173405

Ehrlich, S., & Maestas, C. (2010). Risk orientation, risk exposure, and policy opinions: The case of free trade. *Political Psychology*, *31*(5), 657–684.

Eidelman, S., Crandall, C. S., Goodman, J. A., & Blanchar, J. C. (2012). Low-effort thought promotes political conservatism. *Personality & Social Psychology Bulletin*, *38*(6), 808–820. https://doi.org/10.1177/0146167212439213

Ellis, C., & Stimson, J. A. (2009). Symbolic ideology in the American electorate. *Electoral Studies*, *28*(3), 388–402. https://doi.org/10.1016/j.electstud.2009.05.010

Evans, G., Heath, A., & Lalljee, M. (1996). Measuring left-right and libertarian-authoritarian values in the British electorate. *British Journal of Sociology*, *47*(1), 93–112.

Federico, C. M., Hunt, C. V., & Ergun, D. (2009). Political Expertise, Social Worldviews, and Ideology: Translating "Competitive Jungles" and "Dangerous Worlds" into Ideological Reality. *Social Justice Research*, *22*(2–3), 259–279. https://doi.org/10.1007/s11211-009-0097-0

Federico, C. M., & Malka, A. (2018). The Contingent, Contextual Nature of the Relationship Between Needs for Security and Certainty and Political Preferences: Evidence and Implications. *Political Psychology*, *39*(S1), 3–48. https://doi.org/10.1111/pops.12477

*Federico, C. M., Weber, C. R., Ergun, D., & Hunt, C. (2013). Mapping the Connections between Politics and Morality: The Multiple Sociopolitical Orientations Involved in Moral Intuition. *Political Psychology*, *34*(4), 589–610. https://doi.org/10.1111/pops.12006

Feinberg, M., Antonenko, O., Willer, R., Horberg, E. J., & John, O. P. (2014). Gut check: Reappraisal of disgust helps explain liberal–conservative differences on issues of purity. *Emotion*, *14*(3), 513.

*Feinberg, M., & Willer, R. (2015). From Gulf to Bridge: When Do Moral Arguments Facilitate Political Influence? *Personality and Social Psychology Bulletin*. https://doi.org/10.1177/0146167215607842

Feinberg, M., Willer, R., Antonenko, O., & John, O. P. (2012). Liberating Reason From the Passions Overriding Intuitionist Moral Judgments Through Emotion Reappraisal. *Psychological Science*, *23*(7), 788–795. https://doi.org/10.1177/0956797611434747

Feldman, S. (2013). Political ideology. In L. Huddy, D. O. Sears, & J. S. Levy (Eds.), *Oxford Handbook of Political Psychology* (pp. 591–626). New York, NY: Oxford University Press.

Feldman, S., & Johnston, C. (2014). Understanding the Determinants of Political Ideology: Implications of Structural Complexity. *Political Psychology*, *35*(3), 337–358. https://doi.org/10.1111/pops.12055

*Frankowski, S. D. (2014). *Developing and testing an implicit measure of moral foundation accessibility* (Doctoral dissertation). The University of Texas at El Paso, El Paso, TX.

*Frankowski, S. D. (2018). *Unpublished dataset*.

Frimer, J. A., Biesanz, J. C., Walker, L. J., & MacKinlay, C. W. (2013). Liberals and conservatives rely on common moral foundations when making moral judgments about influential people. *Journal of Personality and Social Psychology*, *104*(6), 1040–1059. http://dx.doi.org/10.1037/a0032277

Gelman, A., & Carlin, J. (2014). Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspectives on Psychological Science*, *9*(6), 641–651. https://doi.org/10.1177/1745691614551642

Gerber, A. S., Huber, G. A., Doherty, D., Dowling, C. M., & Ha, S. E. (2010). Personality and Political Attitudes: Relationships across Issue Domains and Political Contexts. *American Political Science Review*, *104*(01), 111–133. https://doi.org/10.1017/S0003055410000031

Giammarco, E. A. (2016). The measurement of individual differences in morality. *Personality and Individual Differences*, *88*, 26–34. https://doi.org/10.1016/j.paid.2015.08.039

Gilligan, C. (1982). *In A Different Voice: Psychological Theory and Women's Development*. Cambridge, MA: Harvard University Press.

Graham, J., Englander, Z., Morris, J. P., Hawkins, C. B., Haidt, J., & Nosek, B. A. (2012). Warning Bell: Liberals Implicitly Respond to Group Morality Before Rejecting it Explicitly. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2071499

Graham, J., & Haidt, J. (2012). Sacred values and evil adversaries: A moral foundations approach. In P. Shaver & M. Mikulincer (Eds.), *The social psychology of morality: Exploring the causes of good and evil* (pp. 11–31). Retrieved from http://www-bcf.usc.edu/~jessegra/papers/Graham&Haidt.in_press.Sacredness.Herzliya_chapter.pdf

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. B., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, *47*, 55–130. https://doi.org/10.1016/B978-0-12-407236-7.00002-4

Graham, J., Haidt, J., Motyl, M., Meindl, P., Iskiwitch, C., & Mooijman, M. (2016). Moral Foundations Theory: On the Advantages of Moral Pluralism Over Moral Monism. In K. Gray & J. Graham (Eds.), *The Atlas of Moral Psychology*. Retrieved from http://www-bcf.usc.edu/~jessegra/papers/GHMMIM.MFT%20Atlas%20chapter.pdf

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*(5), 1029–1046. https://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 & Social Psychology*, *101*(2), 366–385. https://doi.org/10.1037/a0021847

Gray, K., & Keeney, J. E. (2015). Impure or Just Weird? Scenario Sampling Bias Raises Questions About the Foundation of Morality. *Social Psychological and Personality Science*, *6*(8), 859–868. https://doi.org/10.1177/1948550615592241

Gray, K., Young, L., & Waytz, A. (2012). Mind Perception Is the Essence of Morality. *Psychological Inquiry*, *23*(2), 101–124. https://doi.org/10.1080/1047840X.2012.651387

Hadarics, M., & Kende, A. (2018). The Dimensions of Generalized Prejudice within the Dual-Process Model: The Mediating Role of Moral Foundations. *Current Psychology*, *37*(4), 731–739. https://doi.org/10.1007/s12144-016-9544-x

Haidt, J. (2012). *The righteous mind. Why Good People are Divided by Politics and Religion*. New York, NY: Vintage Books/Random House Inc.

*Hannikainen, I. R., Miller, R. M., & Cushman, F. A. (2017). Act Versus Impact: Conservatives and Liberals Exhibit Different Structural Emphases in Moral Judgment. *Ratio, (4)*30. https://doi.org/10.1111/rati.12162

*Harnish, R. J., Bridges, K. R., & Gump, J. T. (2017). Predicting Economic, Social, and Foreign Policy Conservatism: The Role of Right-Wing Authoritarianism, Social Dominance Orientation, Moral Foundations Orientation, and Religious Fundamentalism. *Current Psychology*, 1–12. https://doi.org/10.1007/s12144-016-9552-x

Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect. *Journal of Personality and Social Psychology*, *82*(6), 903–918. http://dx.doi.org/10.1037/0022-3514.82.6.903

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hibbing, J. R., Smith, K. B., & Alford, J. R. (2014). Differences in negativity bias underlie variations in political ideology. *Behavioral and Brain Sciences*, *37*(3), 297–307. https://doi.org/10.1017/S0140525X13001192

*Hirsh, J. B., DeYoung, C. G., Xu, X., & Peterson, J. B. (2010). Compassionate Liberals and Polite Conservatives: Associations of Agreeableness With Political Ideology and Moral Values. *Personality and Social Psychology Bulletin*, *36*(5), 655–664. https://doi.org/10.1177/0146167210366854

*Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday life. *Science (New York, N.Y.)*, *345*(6202), 1340–1343. https://doi.org/10.1126/science.1251560

Inbar, Y., Pizarro, D. A., & Bloom, P. (2009). Conservatives are more easily disgusted than liberals. *Cognition & Emotion*, *23*(4), 714–725. https://doi.org/10.1080/02699930802110007

Inglehart, R., & Klingemann, H. -D. (1976). Party identification, ideological preference and the left-right dimensions among Western mass publics. In I. Budge, I. Crewe & D. Farlie (Eds.), Party identification and beyond (pp. 243–273). London, UK: Wiley.

Inglehart, R., & Welzel, C. (2010). Changing mass priorities: The link between modernization and democracy. *Perspectives on Politics*, *8*(02), 551–567.

Iyer, R., Koleva, S., Graham, J., Ditto, P., & Haidt, J. (2012). Understanding Libertarian Morality: The Psychological Dispositions of Self-Identified Libertarians. *PLoS ONE*, *7*(8), e42366. https://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*(3), 219–236. https://doi.org/10.1177/1088868313480274

Janoff-Bulman, R., & Carnes, N. C. (2014). Motivation and morality: Insights into political ideology. *Behavioral and Brain Sciences*, *37*(3), 316–317. https://doi.org/10.1017/S0140525X13002604

Janoff-Bulman, R., & Carnes, N. C. (2016). Social Justice and Social Order: Binding Moralities across the Political Spectrum. *PLOS ONE*, *11*(3), e0152479. https://doi.org/10.1371/journal.pone.0152479

Joe, V. C., Jones, R. N., & Ryder, S. (1977). Conservatism, Openness to Experience and Sample Bias. *Journal of Personality Assessment*, *41*(5), 527–531. https://doi.org/10.1207/s15327752jpa4105_11

*Johnson, K. A., Hook, J. N., Davis, D. E., Van Tongeren, D. R., Sandage, S. J., & Crabtree, S. A. (2016). Moral foundation priorities reflect U.S. Christians' individual differences in religiosity. *Personality and Individual Differences*, *100*, 56–61. https://doi.org/10.1016/j.paid.2015.12.037

Jones, B. (2011). Intuitive Politics: How Moral Intuitions Shape Political Identities and Behavior. *APSA 2011 Annual Meeting*. Presented at the Rochester, NY. Retrieved from https://papers.ssrn.com/abstract=1902186

Jost, J. T. (2006). The end of the end of ideology. *American Psychologist*, *61*(7), 651.

Jost, J. T., Federico, C. M., & Napier, J. L. (2009). Political Ideology: Its Structure, Functions, and Elective Affinities. *Annual Review of Psychology*, *60*(1), 307–337. https://doi.org/10.1146/annurev.psych.60.110707.163600

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, *129*(3), 339–375. https://doi.org/10.1037/0033-2909.129.3.339

*Kim, K. R., Kang, J.-S., & Yun, S. (2012). Moral intuitions and political orientation: Similarities and differences between korea and the united states. *Psychological Reports*, *111*(1), 173–185. https://doi.org/10.2466/17.09.21.PR0.111.4.173-185

Kinder, D. R., & Kalmoe, N. P. (2017). *Neither liberal nor conservative: Ideological innocence in the American public*. University of Chicago Press.

Kivikangas, J. M., Lönnqvist, J.-E., & Ravaja, N. (2016). *Relationships between Moral Foundations and political orientation – local study and meta-analysis.* [Poster]. Annual convention of Society for Personality and Social Psychology, San Diego, CA. DOI: 10.13140/RG.2.1.2277.0964

*Kivikangas, J. M., Lönnqvist, J.-E., & Ravaja, N. (2017). Relationship of Moral Foundations to Political Liberalism-Conservatism and Left-Right Orientation in a Finnish Representative Sample. *Social Psychology*, *48*(4), 246–251. https://doi.org/10.1027/1864-9335/a000297

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams, R. B., Alper, S., … Axt, J. R. (2018). *Many Labs 2: Investigating Variation in Replicability Across Sample and Setting*. Retrieved from https://psyarxiv.com/9654g/

*Kliethermes, B. C. (2015). *Perceptions of Computer-Generated Child Pornography* (M.A., The University of North Dakota). Retrieved from https://search.proquest.com/docview/1760170258/abstract/39558BD0CA52423APQ/1

Kohlberg, L. (1971). From is to ought: How to commit the naturalistic fallacy and get away with it. *Cognitive Development and Epistemology*, 151–232.

Koleva, S., 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. https://doi.org/10.1016/j.jrp.2012.01.006

Koleva, S., Selterman, D., Kang, H., & Graham, J. (2014). Beyond Kohlberg vs. Gilligan: Empathy and Disgust Sensitivity Mediate Gender Differences in Moral Judgments. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2486030

Kossowska, M., & Hiel, A. V. (2003). The relationship between need for closure and conservative beliefs in Western and Eastern Europe. Political Psychology, 24(3), 501-518. https://doi.org/10.1111/0162-895X.00338

Koster, W. de, Achterberg, P., & Waal, J. van der. (2013). The new right and the welfare state: The electoral relevance of welfare chauvinism and welfare populism in the Netherlands. *International Political Science Review*, *34*(1), 3–20. https://doi.org/10.1177/0192512112455443

Kromrey, J. D., & Rendina-Gobioff, G. (2006). On Knowing What We Do Not Know: An Empirical Comparison of Methods to Detect Publication Bias in Meta-Analysis. *Educational and Psychological Measurement*, *66*(3), 357–373. https://doi.org/10.1177/0013164405278585

*Kugler, M., Jost, J. T., & Noorbaloochi, S. (2014). Another Look at Moral Foundations Theory: Do Authoritarianism and Social Dominance Orientation Explain Liberal-Conservative Differences in "Moral" Intuitions? *Social Justice Research*, *27*(4), 413–431. https://doi.org/10.1007/s11211-014-0223-5

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(1), 1–26. https://doi.org/10.18637/jss.v082.i13

Lakens, D., & Evers, E. R. (2014). Sailing from the seas of chaos into the corridor of stability practical recommendations to increase the informational value of studies. *Perspectives on Psychological Science*, *9*(3), 278–292.

Lakens, D., Hilgard, J., & Staaks, J. (2016). On the reproducibility of meta-analyses: Six practical recommendations. *BMC Psychology*, *4*, 24. https://doi.org/10.1186/s40359-016-0126-3

*Landy, J. F. (2016). Representations of moral violations: Category members and associated features. *Judgment and Decision Making; Tallahassee*, *11*(5), 496–508.

*Lewis, G. J., & Bates, T. C. (2011). From left to right: How the personality system allows basic traits to influence politics via characteristic moral adaptations. *British Journal of Psychology*, *102*(3), 546–558. https://doi.org/10.1111/j.2044-8295.2011.02016.x

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., … Moher, D. (2009). The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *PLOS Medicine*, *6*(7), e1000100. https://doi.org/10.1371/journal.pmed.1000100

Lilla, M. (2016). *The Shipwrecked Mind: On Political Reaction*. New York Review of Books.

Macaskill, P., Walter, S. D., & Irwig, L. (2001). A comparison of methods to detect publication bias in meta-analysis. *Statistics in Medicine*, *20*(4), 641–654. https://doi.org/10.1002/sim.698

Malka, A., Lelkes, Y., & Soto, C. J. (2019). Are cultural and economic conservatism positively correlated? A large-scale cross-national test. British Journal of Political Science, 49(3), 1045-1069.

Malka, A., Osborne, D., Soto, C. J., Greaves, L. M., Sibley, C. G., & Lelkes, Y. (2016). Binding Moral Foundations and the Narrowing of Ideological Conflict to the Traditional Morality Domain. *Personality and Social Psychology Bulletin*, *42*(9), 1243–1257. https://doi.org/10.1177/0146167216653936

Malka, A., Soto, C. J., Inzlicht, M., & Lelkes, Y. (2014). Do needs for security and certainty predict cultural and economic conservatism? A cross-national analysis. *Journal of Personality and Social Psychology*, *106*(6), 1031. https://doi.org/10.1037/a0036170

*Marszalek, J. M., Barber, C., & Nilsson, J. E. (2017). Development of the Social Issues Advocacy Scale-2 (SIAS-2). *Social Justice Research*, *30*(2), 117–144. https://doi.org/10.1007/s11211-017-0284-3

McFarland, S. (2010). Authoritarianism, social dominance, and other roots of generalized prejudice. *Political Psychology*, *31*(3), 453–477.

*Milesi, P. (2017). Moral foundations and voting intention in Italy. *Europe's Journal of Psychology*, *13*(4), 667–687. https://doi.org/10.5964/ejop.v13i4.1391

*Milesi, P., & Alberici, A. I. (2018). Pluralistic morality and collective action: The role of moral foundations. *Group Processes & Intergroup Relations*, *21*(2), 235–256. https://doi.org/10.1177/1368430216675707

Milojev, P., Osborne, D., Greaves, L. M., Bulbulia, J., Wilson, M. S., Davies, C. L., … Sibley, C. G. (2014). Right-wing authoritarianism and social dominance orientation predict different moral signatures. *Social Justice Research*, *27*(2), 149–174.

Moeyaert, M., Ugille, M., Natasha Beretvas, S., Ferron, J., Bunuan, R., & Van den Noortgate, W. (2017). Methods for dealing with multiple outcomes in meta-analysis: A comparison between averaging effect sizes, robust variance estimation and multilevel meta-analysis. *International Journal of Social Research Methodology*, *20*(6), 559–572.

Moher, D., Jadad, A. R., & Tugwell, P. (1996). Assessing the Quality of Randomized Controlled Trials: Current Issues and Future Directions. *International Journal of Technology Assessment in Health Care*, *12*(2), 195–208. https://doi.org/10.1017/S0266462300009570

Mõttus, R. (2016). Towards More Rigorous Personality Trait–Outcome Research. *European Journal of Personality*, *30*(4), 292–303. https://doi.org/10.1002/per.2041

Napier, J. L., & Luguri, J. B. (2013). Moral Mind-Sets: Abstract Thinking Increases a Preference for "Individualizing" Over "Binding" Moral Foundations. *Social Psychological and Personality Science*, *4*(6), 754–759. https://doi.org/10.1177/1948550612473783

Necka, E. A., Cacioppo, S., Norman, G. J., & Cacioppo, J. T. (2016). Measuring the Prevalence of Problematic Respondent Behaviors among MTurk, Campus, and Community Participants. *PLOS ONE*, *11*(6), e0157732. https://doi.org/10.1371/journal.pone.0157732

*New Zealand Attitudes and Values Study. (2012). *NZAVS Wave 3.5*. Retrieved from https://www.psych.auckland.ac.nz/en/about/our-research/research-groups/new-zealand-attitudes-and-values-study.html

*Niemi, L., & Young, L. (2013). Caring across Boundaries versus Keeping Boundaries Intact: Links between Moral Values and Interpersonal Orientations. *PLOS ONE*, *8*(12), e81605. https://doi.org/10.1371/journal.pone.0081605

Niemi, L., & Young, L. (2016). When and why we see victims as responsible: The impact of ideology on attitudes toward victims. *Personality and Social Psychology Bulletin*, *42*(9), 1227–1242.

*Nilsson, A., & Erlandsson, A. (2015). The Moral Foundations taxonomy: Structural validity and relation to political ideology in Sweden. *Personality and Individual Differences*, *76*, 28–32. https://doi.org/10.1016/j.paid.2014.11.049

*Nilsson, A., Erlandsson, A., & Västfjäll, D. (2016). The congruency between moral foundations and intentions to donate, self-reported donations, and actual donations to charity. *Journal of Research in Personality*, *65*, 22–29. https://doi.org/10.1016/j.jrp.2016.07.001

*Nilsson, A., & Strupp-Levitsky, M. (2016). Humanistic and normativistic metaphysics, epistemology, and conative orientation: Two fundamental systems of meaning. *Personality and Individual Differences*, *100*, 85–94. https://doi.org/10.1016/j.paid.2016.01.050

Norris, C. J., Larsen, J. T., Crawford, L. E., & Cacioppo, J. T. (2011). Better (or worse) for some than others: Individual differences in the positivity offset and negativity bias. *Journal of Research in Personality*, *45*(1), 100–111. https://doi.org/10.1016/j.jrp.2010.12.001

*Okimoto, T. G., & Gromet, D. M. (2016). Differences in sensitivity to deviance partly explain ideological divides in social policy support. *Journal of Personality and Social Psychology*, *111*(1), 98–117. https://doi.org/10.1037/pspp0000080

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Pan, J., & Xu, Y. (2017). *China's Ideological Spectrum* (SSRN Scholarly Paper No. ID 2593377). Retrieved from Social Science Research Network website: https://papers.ssrn.com/abstract=2593377

Pashler, H., & Wagenmakers, E. (2012). Editors' Introduction to the Special Section on Replicability in Psychological Science: A Crisis of Confidence? *Perspectives on Psychological Science*, *7*(6), 528–530. https://doi.org/10.1177/1745691612465253

Philpot, T. (2017). Conservative but not Republican: The paradox of party identification and ideology among African Americans. New York, NY: Cambridge University Press.

Piazza, J., & Landy, J. F. (2013). " Lean not on your own understanding": Belief that morality is founded on divine authority and non-utilitarian moral judgments. *Judgment and Decision Making*, *8*(6), 639.Pinheiro, J. F. M., & Hannikainen, I. R. (2018). *Unpublished dataset*.

Piurko, Y., Schwartz, S. H., & Davidov, E. (2011). Basic Personal Values and the Meaning of Left-Right Political Orientations in 20 Countries. *Political Psychology*, *32*(4), 537–561. https://doi.org/10.1111/j.1467-9221.2011.00828.x

Pollock, P. H., Lilie, S. A., & Vittes, M. E. (1993). Hard Issues, Core Values and Vertical Constraint: The Case of Nuclear Power. *British Journal of Political Science*, *23*(1), 29–50. https://doi.org/10.1017/S0007123400006554

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, *67*(4), 741–763. https://doi.org/10.1037/0022-3514.67.4.741

*Precopio, R. F., & Ramsey, L. R. (2017). Dude looks like a feminist!: Moral concerns and feminism among men. *Psychology of Men & Masculinity*, *18*(1), 78.

R Core Team. (2015). *R: A language and environment for statistical computing*. Retrieved from http://www.R-project.org/

*Regts, G. J. (2015). *Elections Have Consequences: Moral Value Foundations Ensure Gridlock through the Ballot Box*.

Riley, R. D., Lambert, P. C., & Abo-Zaid, G. (2010). Meta-analysis of individual participant data: Rationale, conduct, and reporting. *BMJ*, *340*, c221. https://doi.org/10.1136/bmj.c221

*Ritter, R. S. (2014). *Disgust and religious morality of mind* (Doctoral dissertation). University of Illinois at Urbana-Champaign.

Robin, C. (2011). *The Reactionary Mind: Conservatism from Edmund Burke to Sarah Palin*. Oxford University Press.

*Rodriguez, C., & Rengifo, M. J. (n.d.). *Moral foundations and aggressiveness*. Manuscript in preparation.

*Rossen, I. L., Dunlop, P. D., & Lawrence, C. M. (2015). The desire to maintain the social order and the right to economic freedom: Two distinct moral pathways to climate change scepticism. *Journal of Environmental Psychology*, *42*, 42–47. https://doi.org/10.1016/j.jenvp.2015.01.006

Royzman, E. B., Atanasov, P., Landy, J. F., Parks, A., & Gepty, A. (2014). CAD or MAD? Anger (not disgust) as the predominant response to pathogen-free violations of the divinity code. *Emotion*, *14*(5), 892.

Russell, P. S., & Giner-Sorolla, R. (2013). Bodily moral disgust: What it is, how it is different from anger, and why it is an unreasoned emotion. *Psychological Bulletin, 139*(2), 328–351.

Sanderson, S., Tatt, I. D., & Higgins, J. P. (2007). Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: A systematic review and annotated bibliography. *International Journal of Epidemiology*, *36*(3), 666–676. https://doi.org/10.1093/ije/dym018

Schein, C., & Gray, K. (2015). The Unifying Moral Dyad. Liberals and Conservatives Share the Same Harm-Based Moral Template. *Personality and Social Psychology Bulletin*, *41*(8), 1147–1163. https://doi.org/10.1177/0146167215591501

Schein, C., & Gray, K. (2018). The Theory of Dyadic Morality: Reinventing Moral Judgment by Redefining Harm. *Personality and Social Psychology Review*, *22*(1), 32–70. https://doi.org/10.1177/1088868317698288

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In *Advances in experimental social psychology, Vol. 25* (pp. 1–65). San Diego, CA, US: Academic Press.

Schwartz, S. H. (2007). Universalism values and the inclusiveness of our moral universe. *Journal of Cross-Cultural Psychology*, *38*(6), 711–728.

*Scott, I. M., & Pound, N. (2015). Menstrual Cycle Phase Does Not Predict Political Conservatism. *PLoS ONE*, *10*(4), e0112042. https://doi.org/10.1371/journal.pone.0112042

Seeboth, A., & Mõttus, R. (2018). Successful explanations start with accurate descriptions: Questionnaire items as personality markers for more accurate predictions. *European Journal of Personality*.

Sibley, C. G., & Duckitt, J. (2009). Big-five personality, social worldviews, and ideological attitudes: Further tests of a dual process cognitive-motivational model. *The Journal of Social Psychology*, *149*(5), 545–561.

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York, NY: Cambridge University Press.

*Silver, E., & Abell, L. (2016). Beyond Harm and Fairness: A Study of Deviance and Morality. *Deviant Behavior*, *37*(5), 496–508. https://doi.org/10.1080/01639625.2015.1060746

*Silver, J. R. (2017). Moral Foundations, Intuitions of Justice, and the Intricacies of Punitive Sentiment. *Law & Society Review*, *51*(2), 413–450. https://doi.org/10.1111/lasr.12264

*Simpson, A., Laham, S., & Fiske, A. P. (2016). Wrongness in Different Relationships: Relational Context Effects on Moral Judgment. *The Journal of Social Psychology*, (just-accepted). Retrieved from http://www.tandfonline.com/doi/abs/10.1080/00224545.2016.1140118

*Simpson, A., Piazza, J., & Rios, K. (2016). Belief in divine moral authority: Validation of a shortened scale with implications for social attitudes and moral cognition. *Personality and Individual Differences*, *94*, 256–265. https://doi.org/10.1016/j.paid.2016.01.032

Smith, K. B., Alford, J. R., Hibbing, J. R., Martin, N. G., & Hatemi, P. K. (2016). Intuitive Ethics and Political Orientations: Testing Moral Foundations as a Theory of Political Ideology. *American Journal of Political Science*, *(early view)*. https://doi.org/10.1111/ajps.12255

*Steiger, R. L., & Reyna, C. (2017). Trait contempt, anger, disgust, and moral foundation values. *Personality and Individual Differences*, *113*, 125–135. https://doi.org/10.1016/j.paid.2017.02.071

Sterne, J. A. C., & Egger, M. (2001). Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology*, *54*(10), 1046–1055. https://doi.org/10.1016/S0895-4356(01)00377-8

Sterne, J. A. C., Gavaghan, D., & Egger, M. (2000). Publication and related bias in meta-analysis: Power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology*, *53*(11), 1119–1129.

*Stevens, S. T. (2013). *Political ideology as moral vigilance.* Doctoral dissertation, Rutgers The State University of New Jersey - New Brunswick. Retrieved from https://search.proquest.com/docview/1478020993/abstract/CEB2C86E7DA4427FPQ/1

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

*Teaff, S. H. (2014). *Authentic leadership: An examination of the relationship between political ideology, moral reasoning, and authentic leadership behaviors* (Doctoral dissertation). Our Lady of the Lake University, San Antonio, TX.

Thorisdottir, H., Jost, J. T., Liviatan, I., & Shrout, P. E. (2007). Psychological Needs and Values Underlying Left-Right Political Orientation: Cross-National Evidence from Eastern and Western Europe. *Public Opinion Quarterly*, *71*(2), 175–203. https://doi.org/10.1093/poq/nfm008

Tipton, E., Pustejovsky, J. E., & Ahmadi, H. (2019). Current practices in meta-regression in psychology, education, and medicine. *Research Synthesis Methods*, *10*(2), 180–194. https://doi.org/10.1002/jrsm.1339

Tipton, E., Pustejovsky, J. E., & Ahmadi, H. (2019). A history of meta-regression: Technical, conceptual, and practical developments between 1974 and 2018. *Research Synthesis Methods, 10*, 161–179.

Treier, S., & Hillygus, D. S. (2009). The Nature of Political Ideology in the Contemporary Electorate. *Public Opinion Quarterly*, *73*(4), 679–703. https://doi.org/10.1093/poq/nfp067

Truett, K. R. (1993). Age differences in conservatism. *Personality and Individual Differences*, *14*(3), 405–411. https://doi.org/10.1016/0191-8869(93)90309-Q

*Trups-Kalne, I., & Dimdins, G. (2017). Relation between Social Conservatism, Moral Competence, Moral Orientations, and the Importance of Moral Foundations. *Mediterranean Journal of Social Sciences*, *8*(6), 81–90. https://doi.org/10.1515/mjss-2017-0044

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

*Vainio, A., & Mäkiniemi, J.-P. (2016). How Are Moral Foundations Associated with Climate-Friendly Consumption? *Journal of Agricultural and Environmental Ethics*, *29*(2), 265–283. https://doi.org/10.1007/s10806-016-9601-3

van Aert, R. C. M. (2018). *Meta-Analysis: Shortcomings and potential* (Doctoral dissertation, Tilburg University). https://doi.org/10.31222/osf.io/eqhjd

van Assen, M. A. L. M., van Aert, R. C. M., & Wicherts, J. M. (2015). Meta-analysis using effect size distributions of only statistically significant studies. *Psychological Methods*, *20*(3), 293–309. https://doi.org/10.1037/met0000025

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods*, *45*(2), 576–594. https://doi.org/10.3758/s13428-012-0261-6

van der Linden, S., & Panagopoulos, C. (2019). The O'Reilly factor: An ideological bias in judgments about sexual harassment. *Personality and Individual Differences*, *139*, 198–201. https://doi.org/10.1016/j.paid.2018.11.022

van Leeuwen, F., & Park, J. H. (2009). Perceptions of social dangers, moral foundations, and political orientation. *Personality and Individual Differences*, *47*(3), 169–173. https://doi.org/10.1016/j.paid.2009.02.017

Vevea, J. L., & Hedges, L. V. (1995). A general linear model for estimating effect size in the presence of publication bias. *Psychometrika*, *60*, 419–435.

Vevea, J. L., & Woods, C. M. (2005). Publication bias in research synthesis: Sensitivity analysis using a priori weight functions. *Psychological Methods*, *10*, 428–443.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48.

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Maas, H. L. J., & Kievit, R. A. (2012). An Agenda for Purely Confirmatory Research. *Perspectives on Psychological Science*, *7*(6), 632–638. https://doi.org/10.1177/1745691612463078

*Weber, C. R., & Federico, C. M. (2013). Moral Foundations and Heterogeneity in Ideological Preferences. *Political Psychology*, *34*(1), 107–126. https://doi.org/10.1111/j.1467-9221.2012.00922.x

*Wester, J., Timpano, K. R., Çek, D., Lieberman, D., Fieldstone, S. C., & Broad, K. (2015). Psychological and social factors associated with wastewater reuse emotional discomfort. *Journal of Environmental Psychology*, *42*, 16–23. https://doi.org/10.1016/j.jenvp.2015.01.003

Westfall, J., & Yarkoni, T. (2016). Statistically Controlling for Confounding Constructs Is Harder than You Think. *PLOS ONE*, *11*(3), e0152719. https://doi.org/10.1371/journal.pone.0152719

Wetherell, G. A., Brandt, M. J., & Reyna, C. (2013). Discrimination Across the Ideological Divide The Role of Value Violations and Abstract Values in Discrimination by Liberals and Conservatives. *Social Psychological and Personality Science*, *4*(6), 658–667. https://doi.org/10.1177/1948550613476096

*White, K. R. (2012). *Identifying the antecedents of moral conviction* (Doctoral dissertation). The University of Texas at El Paso, El Paso, TX.

Wicherts, J. M., Veldkamp, C. L. S., Augusteijn, H. E. M., Bakker, M., Aert, V., M, R. C., … M, M. A. L. (2016). Degrees of Freedom in Planning, Running, Analyzing, and Reporting Psychological Studies: A Checklist to Avoid p-Hacking. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01832

Wright, J. C., & Baril, G. (2011). The role of cognitive resources in determining our moral intuitions: Are we all liberals at heart? *Journal of Experimental Social Psychology*, *47*(5), 1007–1012. https://doi.org/10.1016/j.jesp.2011.03.014

*Yalçındağ, B., Özkan, T., Cesur, S., Yilmaz, O., Tepe, B., Piyale, Z. E., … Sunar, D. (2017). An Investigation of Moral Foundations Theory in Turkey Using Different Measures. *Current Psychology*, 1–18. https://doi.org/10.1007/s12144-017-9618-4

*Yilmaz, O., Harma, M., Bahçekapili, H. G., & Cesur, S. (2016). Validation of the Moral Foundations Questionnaire in Turkey and its relation to cultural schemas of individualism and collectivism. *Personality and Individual Differences*, *99*, 149–154. https://doi.org/10.1016/j.paid.2016.04.090

*Yilmaz, O., & Saribay, S. A. (2017). Activating analytic thinking enhances the value given to individualizing moral foundations. *Cognition*, *165*(Supplement C), 88–96. https://doi.org/10.1016/j.cognition.2017.05.009

*Yılmaz, O., Sarıbay, S. A., Bahçekapılı, H. G., & Harma, M. (2016). Political orientations, ideological self-categorizations, party preferences, and moral foundations of young Turkish voters. *Turkish Studies*, *17*(4), 544–566. https://doi.org/10.1080/14683849.2016.1221312

*YourMorals collaboration. (2016). *YourMorals.org dataset*.

Zakrisson, I. (2005). Construction of a short version of the Right-Wing Authoritarianism (RWA) scale. *Personality and Individual Differences*, *39*(5), 863–872. https://doi.org/10.1016/j.paid.2005.02.026

Zaller, J. R. (1992). *The nature and origins of mass opinion*. Cambridge university press.

Zhang, Y., & Li, S. (2015). Two Measures for Cross-Cultural Research on Morality: Comparison and Revision. *Psychological Reports*, *117*(1), 144–166. https://doi.org/10.2466/08.07.PR0.117c15z5

*Zuber, J. M., & Andersen, M. L. (2016). Moral Foundations Theory: An Exploratory Study of Politics and Decision-Making. *Journal of Leadership, Accountability and Ethics*, *13*(2), 74.

Table 1

Reliabilities, Means, and Standard Deviations of MFQ factors, and minimum and maximum correlations between the factors and the political self-placement measure, using independent raw data

Moral foundation	MFQ30 Cronbach α	*M*	*SD*	minimum MFQ-PSP *r*	maximum MFQ-PSP *r*
Care	.612	3.62	0.83	-.550	.205
Fairness	.617	3.60	0.76	-.456	.260
Loyalty	.674	2.74	0.94	-.250	.630
Authority	.717	2.75	0.98	-.007	.630
Sanctity	.824	2.47	1.19	-.065	.681

Note: n ranges from 22,003 (Fairness) to 20,271 (Sanctity). $k = 48$. For reliability indices, calculations for reliability are based on studies that had used all 30 items ($k = 36$, n ranging from 9,268 for loyalty to 12,216 for sanctity). PSP = political self-placement. MFQ = Moral Foundations Questionnaire.

Table 2

Basic meta-analysis on MFQ-PSP association, using

Moral foundation	PSP *r* [95% *CI*]	[95 % *PI*]	*tau* (*SD*)	I^2
Care	-.146 [-.168, -.123]	[-.320, .029]	0.088	85.46%
Fairness	-.215 [-.239, -.191]	[-.407, -.022]	0.097	88.26%
Loyalty	.288 [.261, .315]	[.065, .511]	0.113	91.85%
Authority	.367 [.345, .390]	[.184, .551]	0.093	89.64%
Sanctity	.372 [.347, .397]	[.163, .581]	0.106	91.79%

Note. $k = 89$. MFQ = Moral Foundations Questionnaire. PSP = political self-placement.

Table 3

MFQ-PSP association estimates by sample type, using all PSP measures

Foundation	representative (*k*=4)	convenience (*k*=57)	MTurk (*k*=27)	YourMorals (*k*=1)
Care	-.129 [-.211, -.047]	-.131 [-.160, -.102]	-.175 [-.214, -.136]	-.332 [-.336, -.329]
	[-.320, 062]	*[-.306, .044]*	*[-.352, .002]*	
Fairness	-.167 [-251, -.082]	-.200 [-.231, -.169]	-.232 [-.274, -.191]	-.420 [-.423, -.417]
	[-.372, .038]	*[-.390, -.011]*	*[-.424, -.041]*	
Loyalty	.244 [.146, .342]	.285 [.252, .318]	.300 [.257, .344]	.445 [.442, .448]
	[.010, .478]	*[.070, .501]*	*[.083, .518]*	
Authority	.378 [.294, .461]	.339 [.310, .367]	.406 [.369, .443]	.521 [.518, .524]
	[.173, .582]	*[.150, .527]*	*[.216, .596]*	
Sanctity	.329 [.223, .434]	.342 [.308, .377]	.414 [.368, .459]	.540 [.537, .543]
	[.069, .589]	*[.102, .583]*	*[.171, .656]*	

Note: The shown intervals are 95 % confidence intervals, which are estimated for independent sample types but calculated from the individual-level data for YourMorals sample. The intervals in italics are prediction intervals (which cannot be calculated for the YM sample, $k = 1$).

Table 4

Means and SDs for demographics (age, gender, education, race, religion) and political interest for all studies that reported them, by sample type

Demographics	representative (*k*=4)	MTurk (*k*=27)	convenience (*k*=57)	YourMorals (*k*=1)
Age	49.09 (4.13)	34.53 (3.97)	29.69 (9.75)	33.5
k for Age	4	22	42	1
Male %	44.26 % (5.09 %)	46.57 % (17.0 %)	34.35 % (13.67 %)	53.73 %
k for gender	4	22	49	1
Education	3.11 (0.23)	3.76 (0.59)	3.53 (0.69)	3.96
k for Educ.	3	9	11	1
White %	86.2 %	77.76 % (3.15 %)	70.11 % (11.76 %)	80.6 %
Black %	5.7 %	7.27 % (1.95 %)	6.7 % (4.12 %)	2.4 %
Hispanic %	4.9 %	5.28 % (2.2 %)	10.87 % (11.34 %)	4.1 %
Other race %	3.2 %	9.68 % (3.33 %)	12.26 % (4.08 %)	12.9 %
k for race	1	14	17	1
Christian %	72.7 %	31.17 % (3.02 %)	57 % (12.9 %)	30.9 %
No religion %	10.9 %	51.75 % (24.28 %)	31.99 % (22.47 %)	13.3 %
Other religion %	16.4 %	19.23 % (11.11 %)	21.6 % (7.56 %)	55.8 %
k for religion	1	6	7	1
Political interest	0.96 (0.29)	1.00	0.85	1.37
k for polit. int.	2	1	1	1

Note: Education on scale from 1 (some high school) to 5 (graduate degree or higher); political interest on scale from 0 (no interest at all) to 2 (very interested in politics). The "other" categories were not tested as moderators in the study-level analyses because of the dependency issues and because as categories they are not well defined.

Table 5
Unstandardized regression coefficients from mixed models, MFQ score predicted by PSP and political interest, separated by sample type

	Representative		Convenience		MTurk		YourMorals	
	b	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*
Care								
Intercept	3.558	0.080	3.979	0.146	3.986	0.077	3.703	0.049
Polint	0.175	0.052	0.128	0.107	0.165	0.043	0.115	0.004
PSP	-0.006	0.021	-0.082	0.044	-0.023	0.021	-0.077	0.003
PSP ×Polint	-0.045	0.015	0.001	0.033	-0.030	0.015	-0.070	0.002
Fairness								
Intercept	3.674	0.104	3.803	0.167	3.724	0.130	3.657	0.120
Polint	0.134	0.044	0.153	0.091	0.130	0.036	0.201	0.004
PSP	-0.047	0.018	-0.086	0.037	-0.018	0.018	-0.096	0.002
PSP ×Polint	-0.015	0.012	-0.012	0.028	-0.029	0.013	-0.066	0.001
Loyalty								
Intercept	2.616	0.224	2.777	0.331	2.169	0.305	1.976	0.298
Polint	-0.026	0.054	-0.061	0.113	-0.083	0.045	-0.069	0.005
PSP	0.153	0.023	0.032	0.046	0.188	0.023	0.255	0.003
PSP ×Polint	0.000	0.015	0.063	0.035	0.024	0.016	-0.003	0.002
Authority								
Intercept	2.515	0.258	3.155	0.377	2.190	0.355	1.979	0.349
Polint	-0.105	0.054	-0.142	0.111	-0.135	0.044	-0.144	0.005
PSP	0.190	0.022	0.010	0.046	0.258	0.022	0.284	0.003
PSP ×Polint	0.001	0.015	0.068	0.035	0.014	0.016	0.015	0.002
Sanctity								
Intercept	2.343	0.348	3.052	0.505	1.887	0.408	1.177	0.475
Polint	-0.101	0.065	-0.187	0.135	-0.171	0.054	-0.134	0.006
PSP	0.174	0.027	-0.039	0.056	0.322	0.027	0.396	0.004
PSP ×Polint	0.017	0.018	0.109	0.042	0.005	0.019	-0.001	0.002

Note. PSP = political self-placement. MFQ = Moral Foundations Questionnaire. Polint = political interest. Political interest measured on scale from 0 to 2.

Table 6
Unstandardized regression coefficients from mixed models, MFQ score predicted by PSP and race, separated by sample type

	Representative		Convenience		MTurk		YM	
	b	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*
Care								
White	3.795	0.246	3.825	0.074	3.818	0.096	3.888	0.241
Other	3.405	0.276	3.688	0.077	3.672	0.104	3.826	0.241
Black	3.385	0.263	3.447	0.080	3.566	0.107	3.744	0.242
Hispanic	3.917	0.265	3.650	0.079	3.662	0.113	3.890	0.242
PSP × White	-0.055	0.013	-0.090	0.006	-0.073	0.008	-0.184	0.001
PSP × Other	-0.023	0.013	-0.058	0.007	-0.041	0.009	-0.152	0.003
PSP × Black	0.023	0.014	-0.012	0.009	0.005	0.011	-0.105	0.007
PSP × Hispanic	0.011	0.014	-0.023	0.008	-0.006	0.010	-0.117	0.005
Fairness								
White	3.182	0.172	3.816	0.061	3.739	0.079	3.960	0.198
Other	3.765	0.228	3.796	0.064	3.757	0.086	3.971	0.198
Black	3.757	0.217	3.670	0.067	3.791	0.088	3.970	0.199
Hispanic	4.024	0.219	3.829	0.065	3.830	0.094	4.024	0.198
PSP × White	0.111	0.014	-0.100	0.005	-0.059	0.007	-0.204	0.001
PSP × Other	-0.026	0.011	-0.070	0.006	-0.029	0.008	-0.174	0.003
PSP × Black	0.027	0.012	-0.018	0.008	0.023	0.009	-0.121	0.006
PSP × Hispanic	-0.007	0.011	-0.052	0.007	-0.010	0.008	-0.155	0.005
Loyalty								
White	2.878	0.137	2.090	0.044	2.079	0.057	1.840	0.127
Other	2.537	0.189	2.243	0.051	2.091	0.071	2.005	0.127
Black	2.777	0.167	2.227	0.056	2.083	0.075	2.073	0.128
Hispanic	3.182	0.172	2.261	0.053	2.132	0.085	2.047	0.128
PSP × White	0.133	0.013	0.230	0.006	0.218	0.009	0.261	0.001
PSP × Other	0.116	0.014	0.214	0.007	0.202	0.009	0.244	0.004
PSP × Black	0.074	0.015	0.171	0.010	0.159	0.011	0.203	0.008
PSP × Hispanic	0.111	0.014	0.209	0.008	0.197	0.010	0.239	0.006
Authority								
White	2.701	0.210	2.068	0.064	1.941	0.083	1.727	0.204
Other	2.781	0.246	2.277	0.068	2.030	0.093	1.830	0.204
Black	3.088	0.230	2.555	0.072	2.496	0.096	2.307	0.205
Hispanic	2.966	0.233	2.319	0.070	2.120	0.104	1.946	0.204
PSP × White	0.163	0.013	0.268	0.006	0.262	0.009	0.319	0.001
PSP × Other	0.145	0.013	0.250	0.007	0.244	0.009	0.301	0.004
PSP × Black	0.101	0.015	0.207	0.009	0.200	0.011	0.258	0.007
PSP × Hispanic	0.141	0.014	0.246	0.008	0.240	0.010	0.297	0.005

Sanctity								
White	2.604	0.276	2.604	0.276	1.557	0.108	0.910	0.269
Other	2.622	0. 316	2.622	0.316	1.675	0.119	1.169	0.269
Black	2.955	0.299	2.954	0.299	2.333	0.123	1.698	0.270
Hispanic	2.888	0.302	2.888	0.302	1.762	0.132	1.182	0.269
PSP × White	0.174	0.016	0.174	0.016	0.335	0.010	0.410	0.001
PSP × Other	0.148	0.016	0.148	0.016	0.309	0.011	0.384	0.004
PSP × Black	0.102	0.018	0.102	0.018	0.263	0.013	0.337	0.009
PSP × Hispanic	0.160	0.017	0.160	0.017	0.321	0.012	0.395	0.007

Note. PSP = political self-placement. MFQ = Moral Foundations Questionnaire.

Table 7

Results of three-level univariate moderator meta-analyses for differences in MFQ-PSP associations between "conservatism" and "right-orientation" labels, by moral foundation

	r [95% *CI*] [95 % *PI*]			
Foundation	conservatism	right-orientation	*z*	*p*
Care	-.129 [-.162, -.095]	-.181 [-.226, -.136]	2.159	.0308
	[-.316., .058]	*[-.370, -.008]*		
Fairness	-.188 [-.225, -.150]	-.228 [-.278, -.178]	1.489	.1364
	[-.401, .026]	*[-.444, -.012]*		
Loyalty	.299 [.258, .340]	.235 [.177, .294]	1.787	.0740
	[.054, .544]	*[-.013, .483]*		
Authority	.366 [.329, .404]	.336 [.282, .389]	0.964	.3352
	[.148, .585]	*[.114, .557]*		
Sanctity	.404 [.366, .442]	.274 [.219, .329]	4.058	< .0001
	[.180, .628]	*[.046, .501]*		

Note. $k = 49$, except for Fairness and Loyalty, k = 48. The intervals in italics are prediction intervals (which cannot be calculated for the YM sample, $k = 1$). PSP = political self-placement. MFQ = Moral Foundations Questionnaire.

Table 8

Results of three-level univariate moderator meta-analyses for differences in MFQ-PSP associations between "social" and "economic" orientation labels, by moral foundation

Foundation	*r* [95% *CI*] [95 % *PI*] social orientation	economic orientation	*z*	*p*
Care	-.138 [-.181, -.096] *[-.284, .007]*	-.178 [-.220, -.136] *[-.324, -.033]*	1.711	.0871
Fairness	-.222 [-.268, -.176] *[-.384, .059]*	-.241 [-.287, -.195] *[-.404, -.078]*	0.919	.3584
Loyalty	.354 [.318, .389] *[.237, .470]*	.251 [.215, .287] *[.134, .368]*	6.299	< .0001
Authority	.431 [.395, .468] *[.308, .555]*	.294 [.256, .331] *[.170, .417]*	8.424	< .0001
Sanctity	.450 [.402, .499] *[.271, .630]*	.250 [.201, .300] *[.071, .430]*	9.112	< .0001

Note. k = 24. The intervals in italics are prediction intervals (which cannot be calculated for the YM sample, k = 1). PSP = political self-placement. MFQ = Moral Foundations Questionnaire.

Table 9

Unstandardized regression coefficients representing the change in MFQ scores per unit of PSP change, for three sample types and a distinction between social and economic orientation.

	Convenience (*b*)		MTurk (*b*)		YourMorals (*b*)	
Foundation	social	economic	social	economic	social	economic
Care	-0.057	-0.053	-0.068	-0.091	-0.111	-0.170
Fairness	-0.077	-0.083	-0.091	-0.107	-0.142	-0.169
Loyalty	0	0.147	0.195	0.143	0.243	0.153
Authority	0	0.151	0.263	0.186	0.301	0.183
Sanctity	0	0.134	0.348	0.191	0.427	0.195

Note. *k* = 14. PSP = political self-placement. MFQ = Moral Foundations Questionnaire.

Table 10

Total effects (unstandardized) of individual-level mixed models for differences between liberalism-conservatism and left-right labels on MFQ-PSP associations, by country.

	Care (*b*)		Fairness (*b*)		Loyalty (*b*)		Authority (*b*)		Sanctity (*b*)	
	Cons	Right	Cons	Right	Cons	Right	Cons	Right	Cons	Right
USA	-0.057	-0.060	-0.050	-0.050	0.186	0.212	0.246	0.259	0.292	0.313
N Zealand	-0.021	-0.058	-0.049	-0.083	0.191	0.180	0.317	0.302	0.404	0.319
Latvia	0.045	-0.015	-0.027	-0.049	0.119	0.053	0.179	-0.051	0.191	-0.045
Finland	-0.046	-0.115	-0.042	-0.113	0.166	0.118	0.226	0.160	0.262	0.093
Sweden	-0.027	-0.074	-0.108	-0.121	0.132	0.110	0.143	0.145	0.158	0.066

Table 11

Results of study-level moderator meta-analyses for differences in MFQ-political orientation associations between political orientations operationalized as RWA and SDO, by moral foundation

Foundation	*r* [95% *CI*] [95 % *PI*] RWA	SDO	*z*	*p*
Care	-.065 [-.130, -.000] *[-.240, .109]*	-.324 [-.390, -.259] *[-.499, -.150]*	7.423	< .0001
Fairness	-.105 [-.178, -.032] *[-.312, .102]*	-.397 [-.470, -.324] *[-.604, -.190]*	6.598	< .0001
Loyalty	.447 [.399, .495] *[.321, .572]*	.216 [.165, .267] *[.090, .342]*	14.499	< .0001
Authority	.553 [.474, .631] *[.322, .784]*	.300 [.218, .382] *[.067, .532]*	5.662	< .0001
Sanctity	.577 [.502, .651] *[.363, .790]*	.186 [.106, .265] *[-.029, .400]*	11.842	< .0001

Note: $k = 10$. The intervals in italics are prediction intervals (which cannot be calculated for the YM sample, $k = 1$).

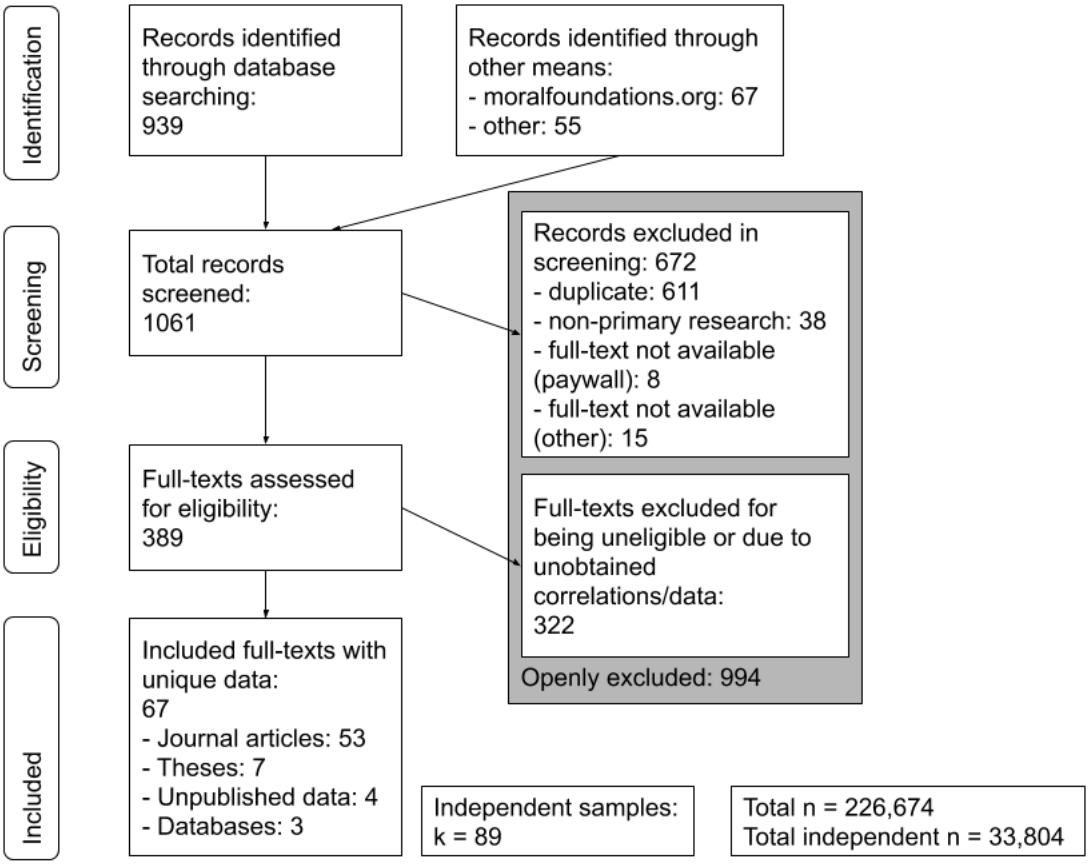

Figure 1. Flowchart of the selection of studies.

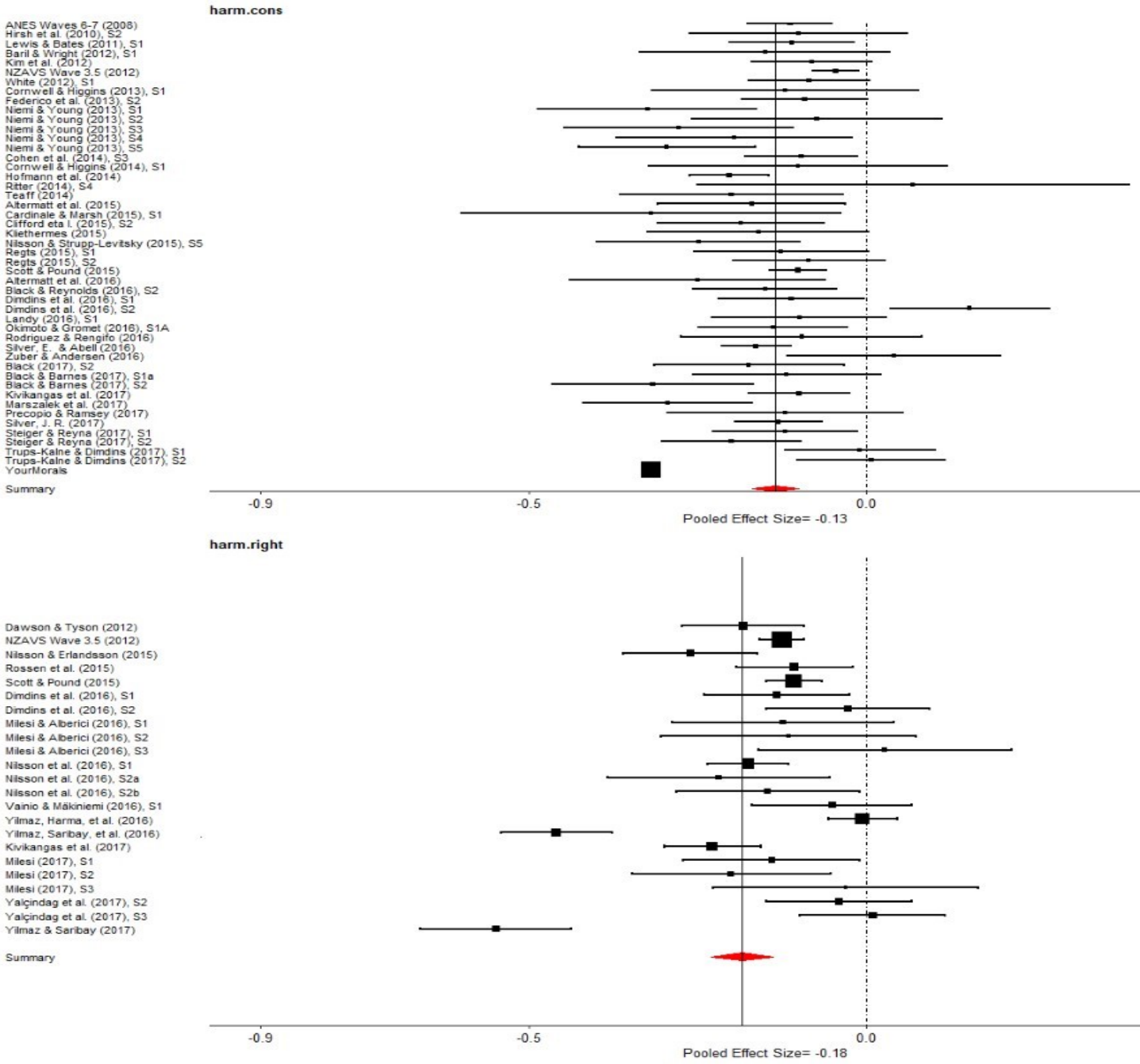

Figure 2. Forest plots for Care-Conservatism and Care-Right.

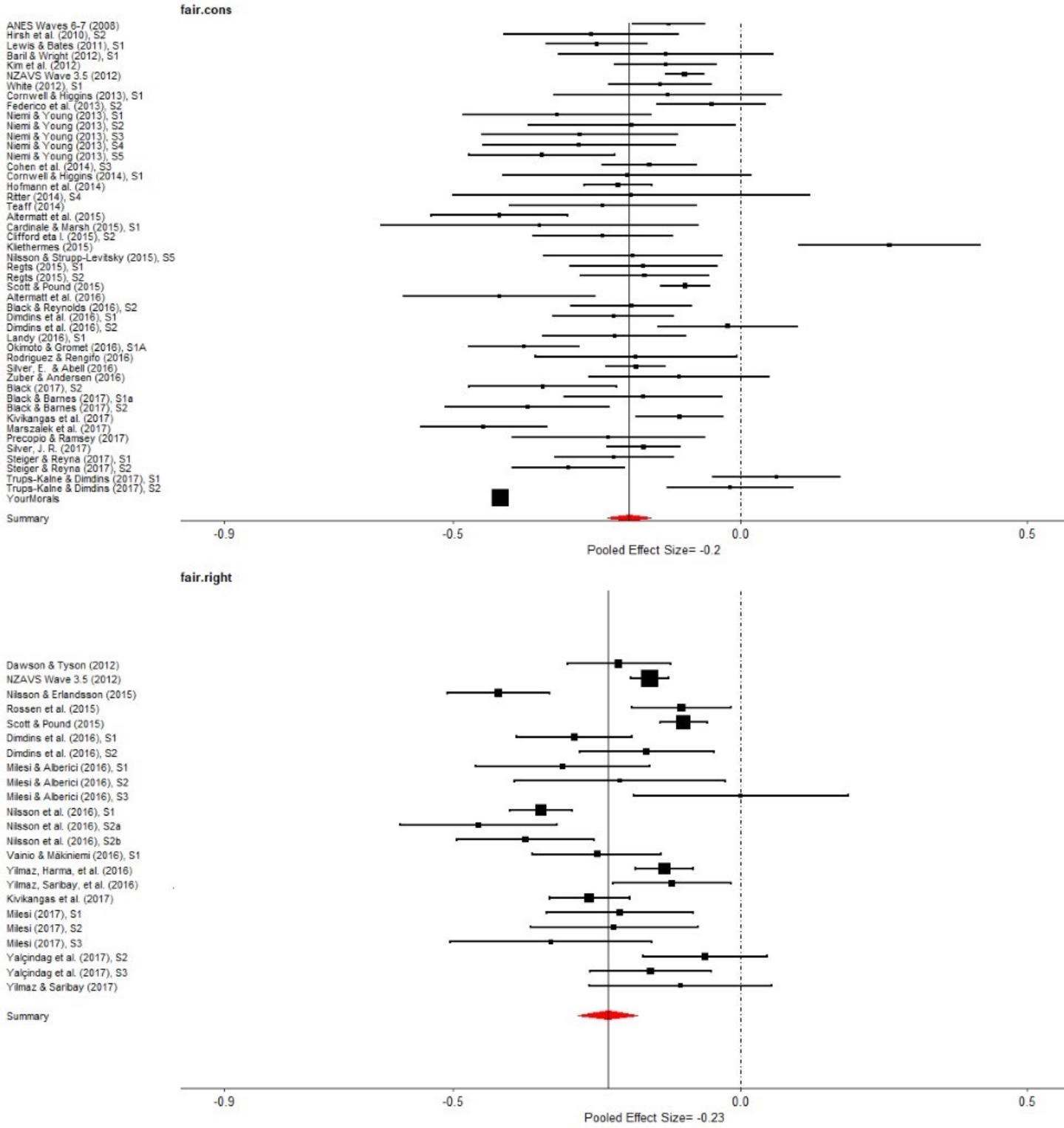

Figure 3. Forest plots for Fairness-Conservatism and Fairness-Right.

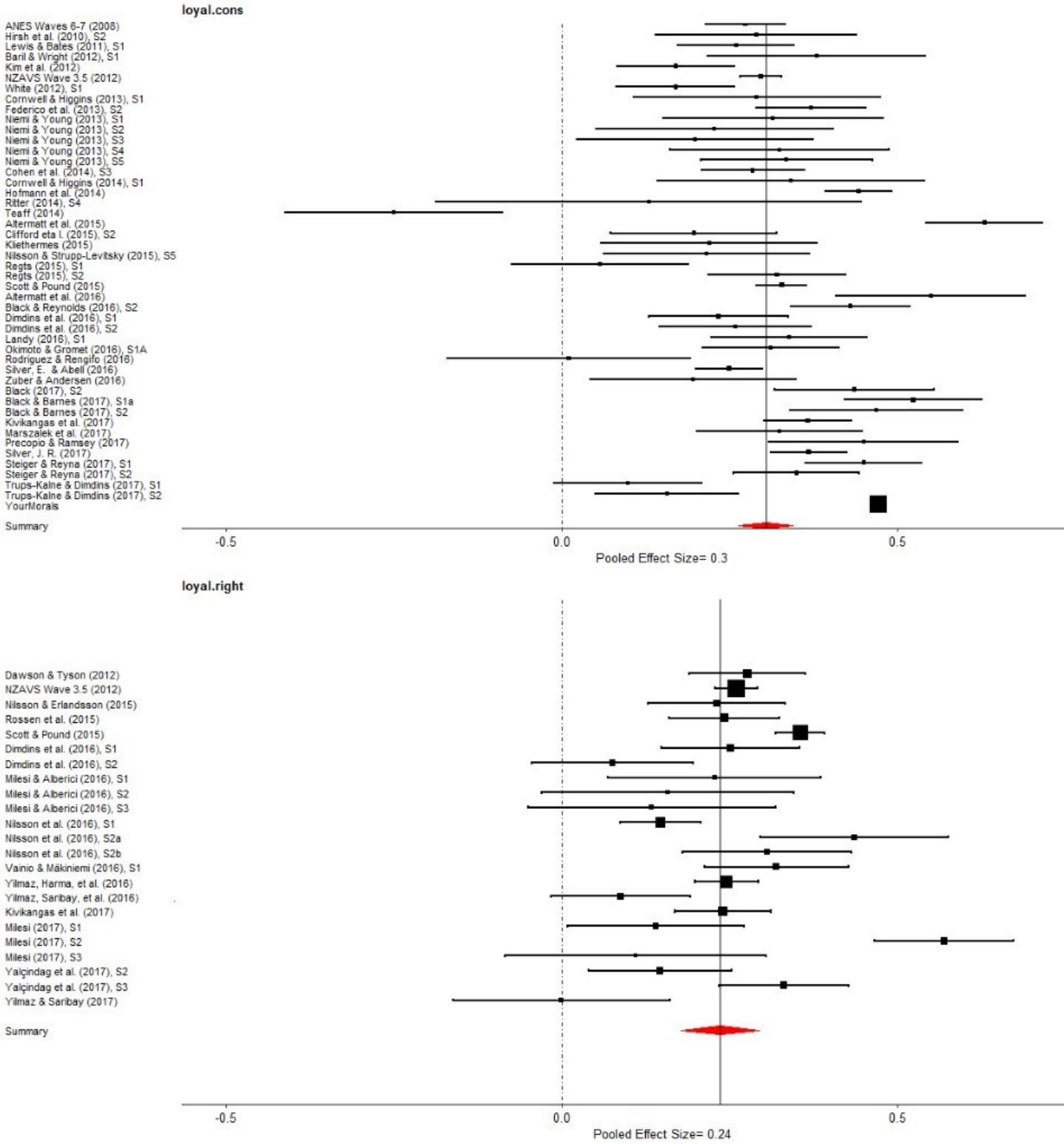

Figure 4. Forest plots for Loyalty-Conservatism and Loyalty-Right.

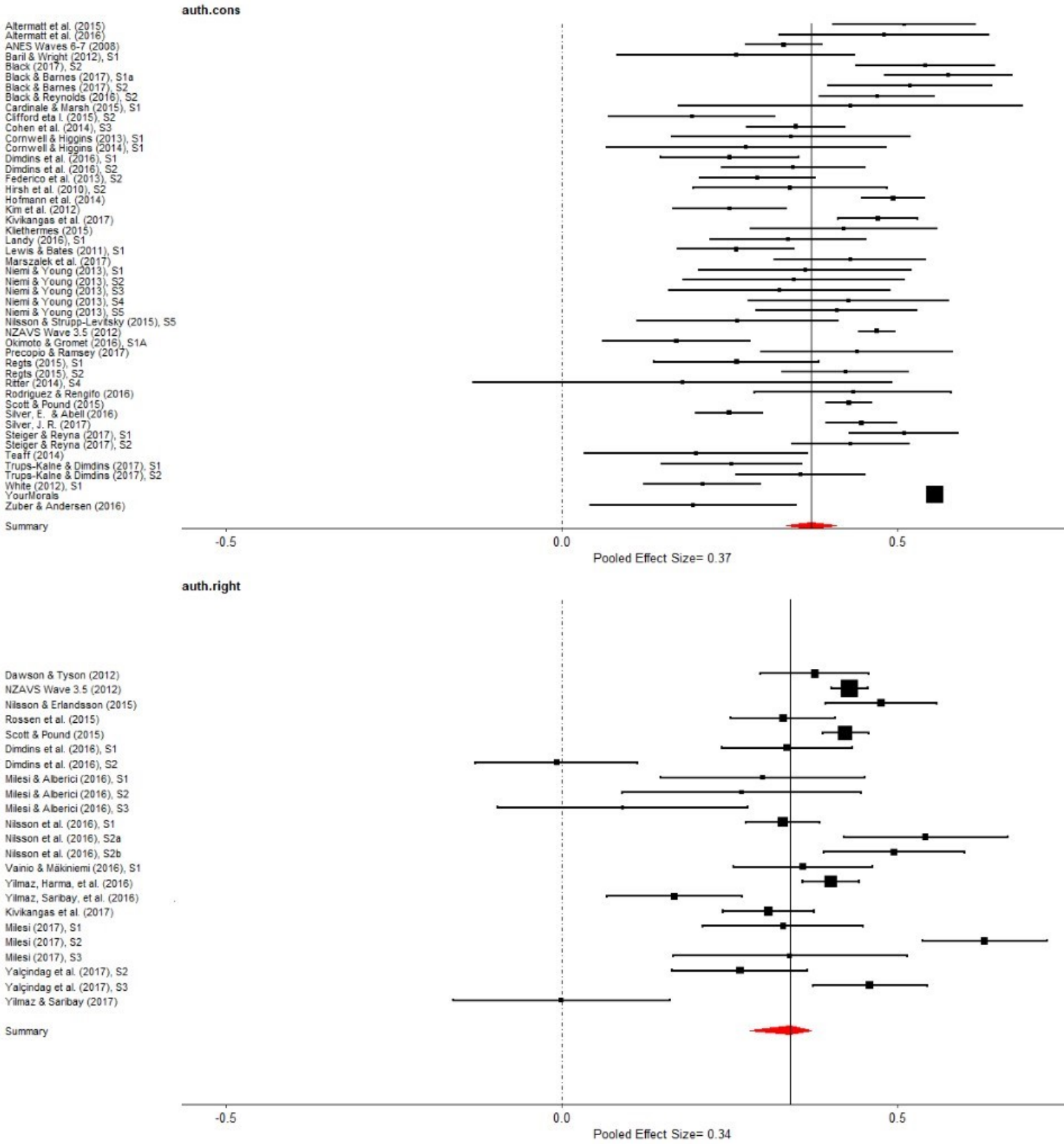

Figure 5. Forest plots for Authority-Conservatism and Authority-Right.

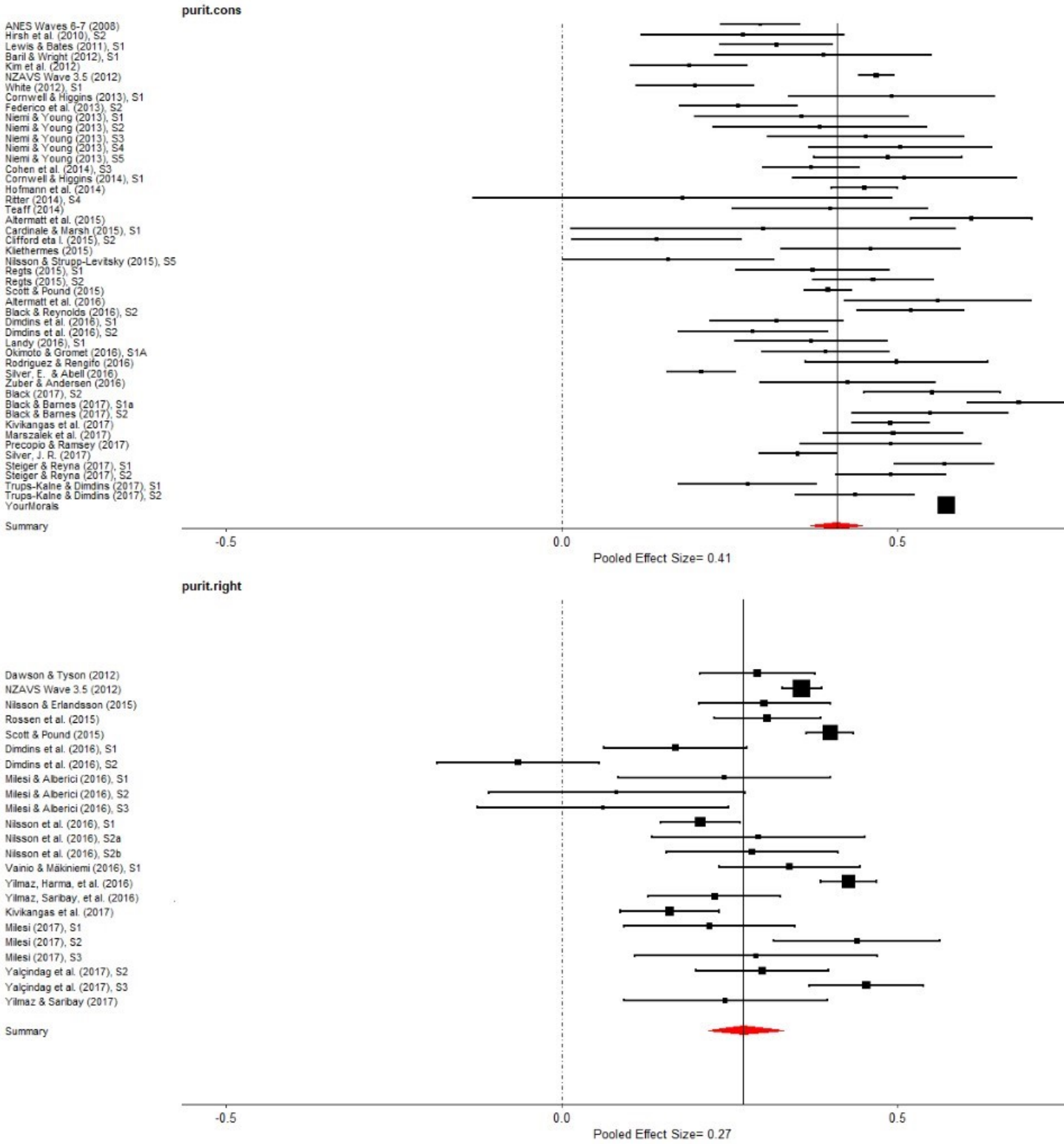

Figure 6. Forest plots for Sanctity-Conservatism and Sanctity-Right.

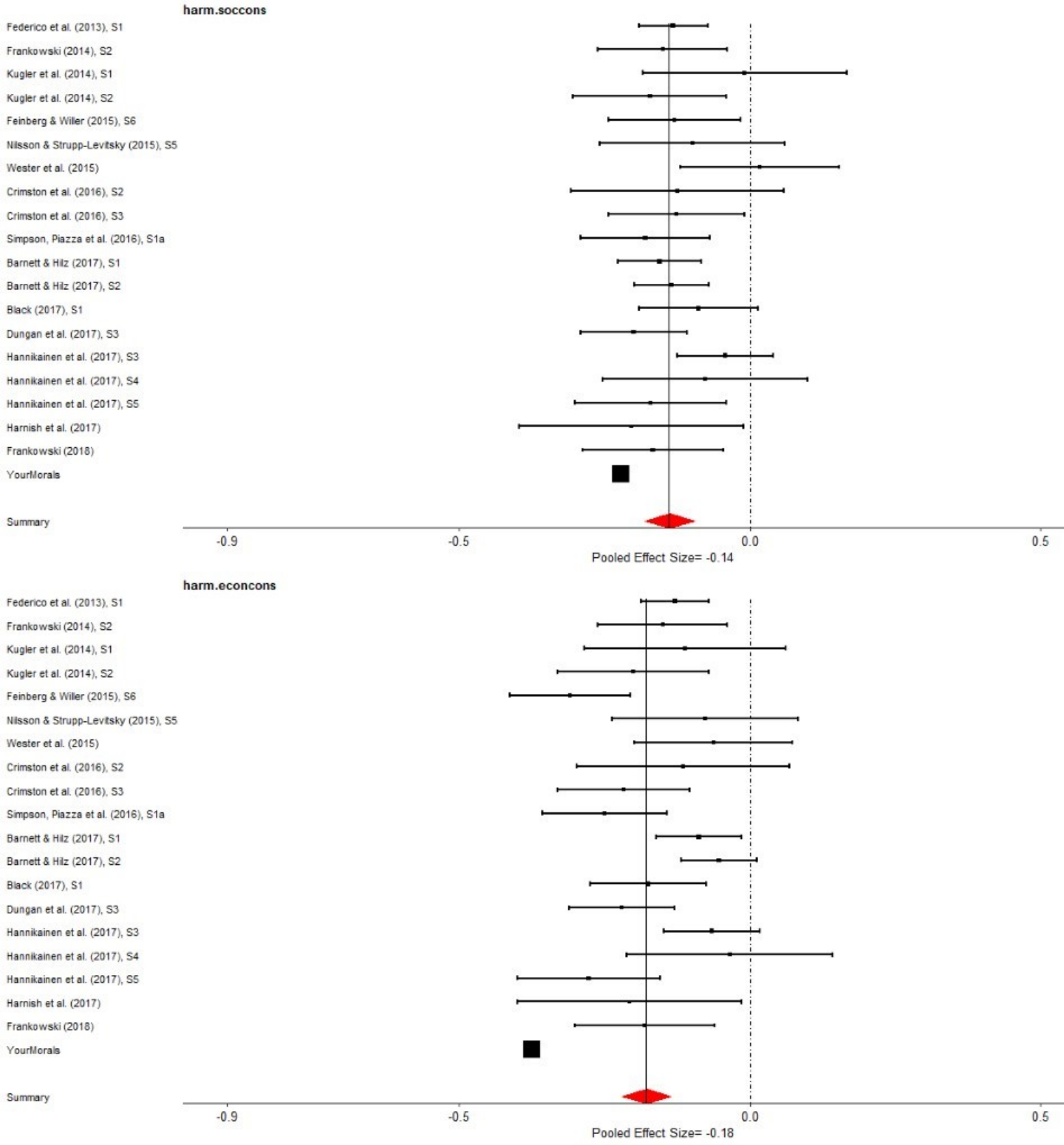

Figure 7. Forest plots for Care-Social conservatism and Care-Economic conservatism.

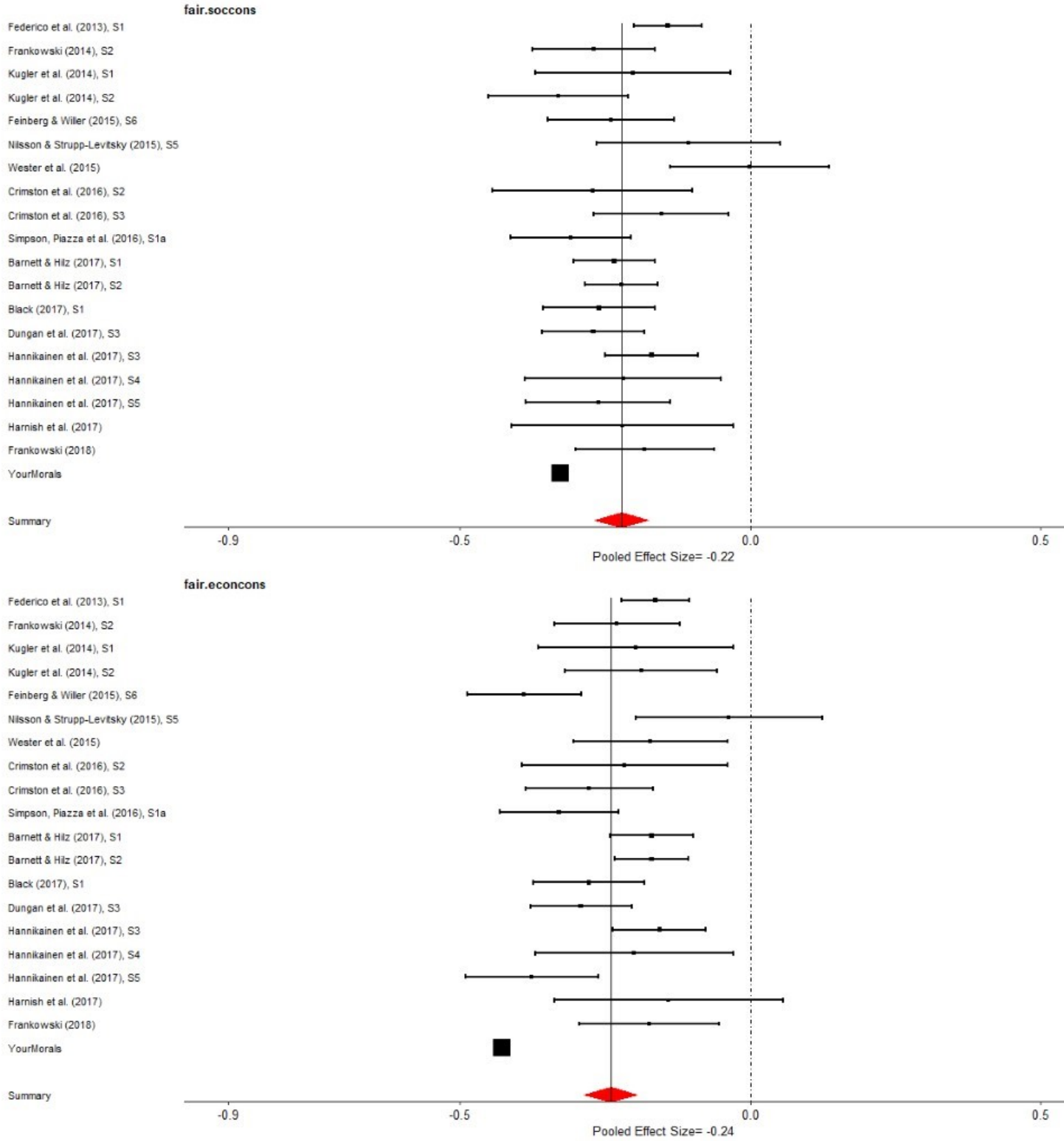

Figure 8. Forest plots for Fairness-Social conservatism and Fairness-Economic conservatism.

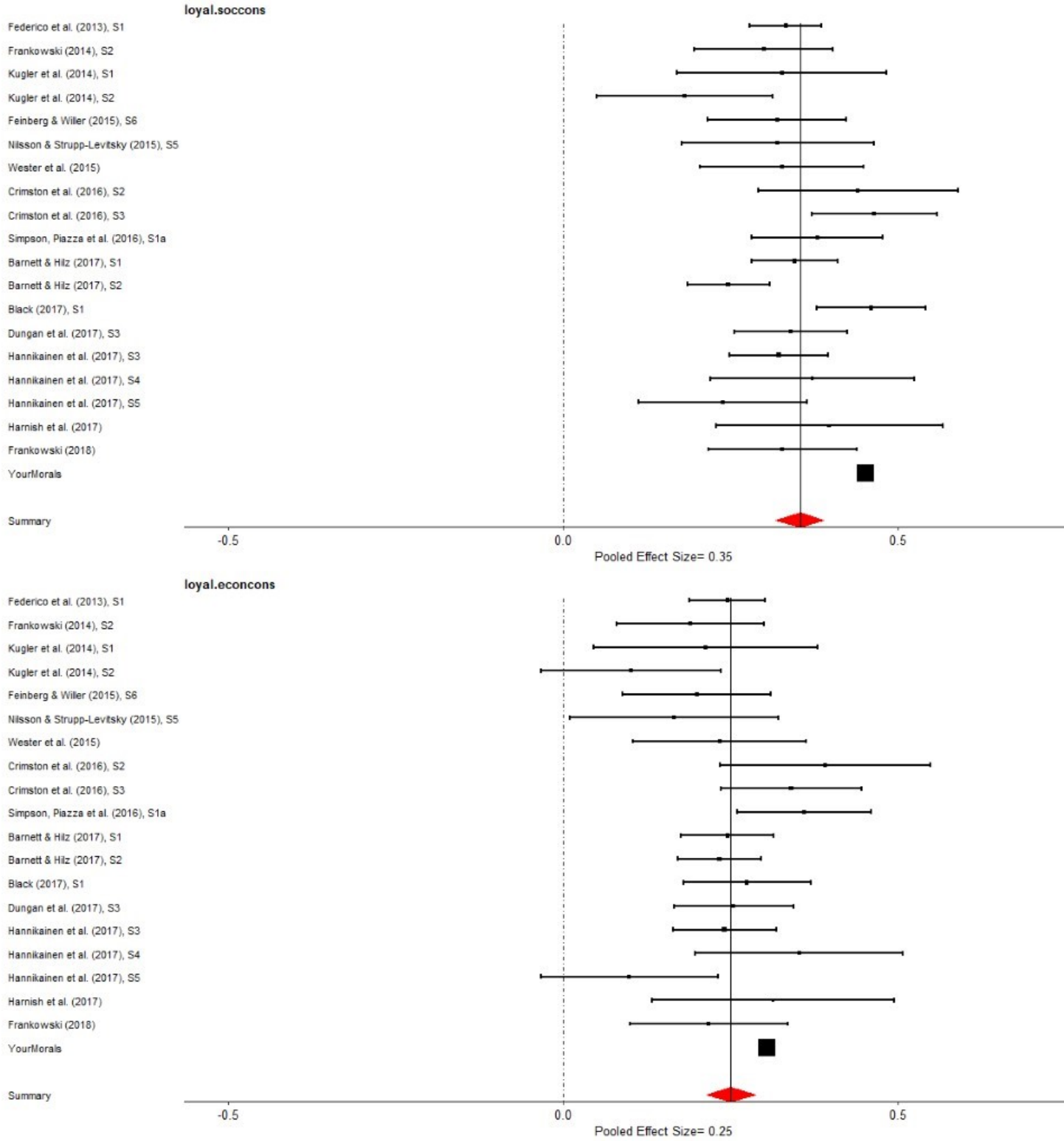

Figure 9. Forest plots for Loyalty-Social conservatism and Loyalty-Economic conservatism.

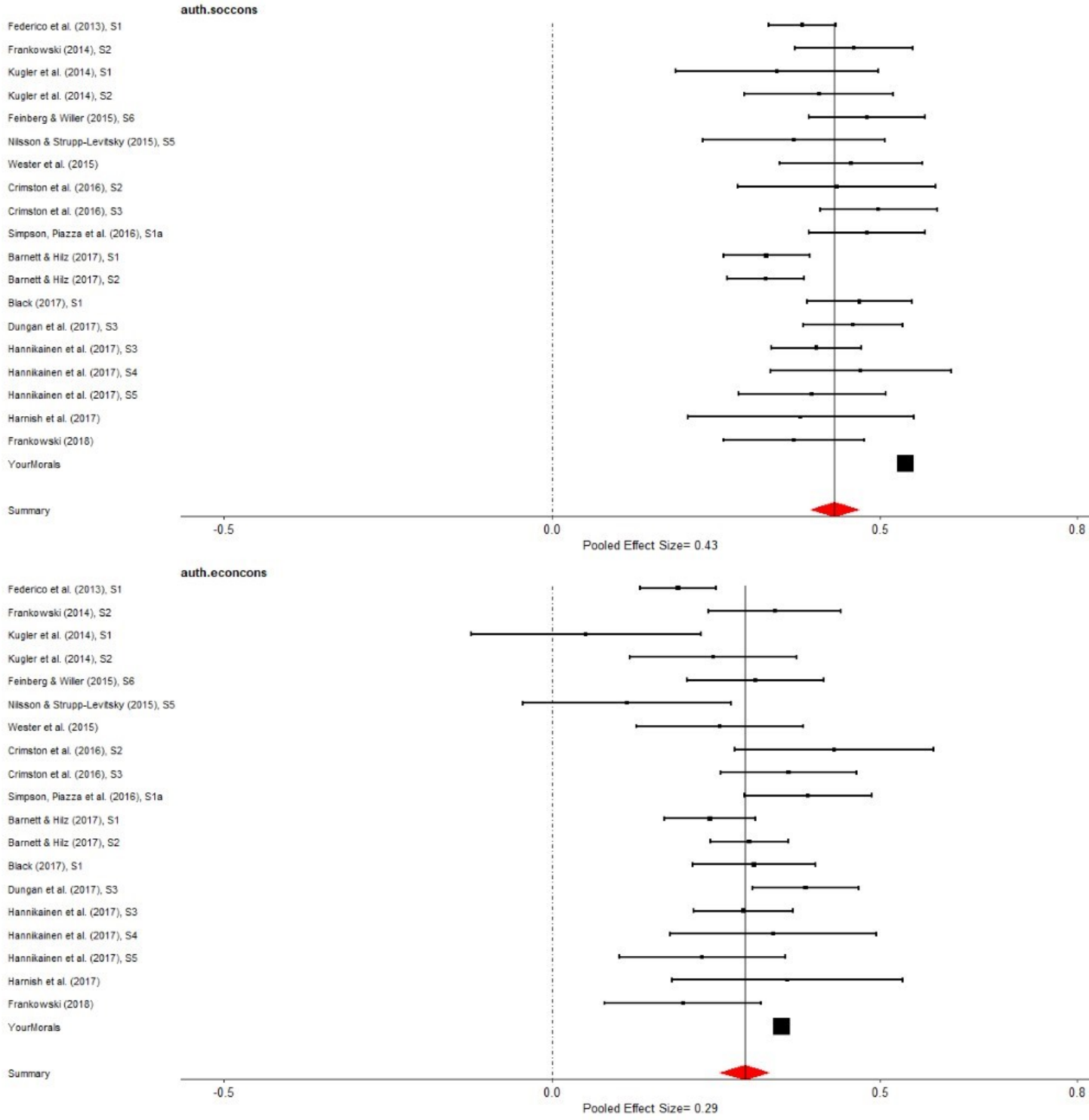

Figure 10. Forest plots for Authority-Social conservatism and Authority-Economic conservatism

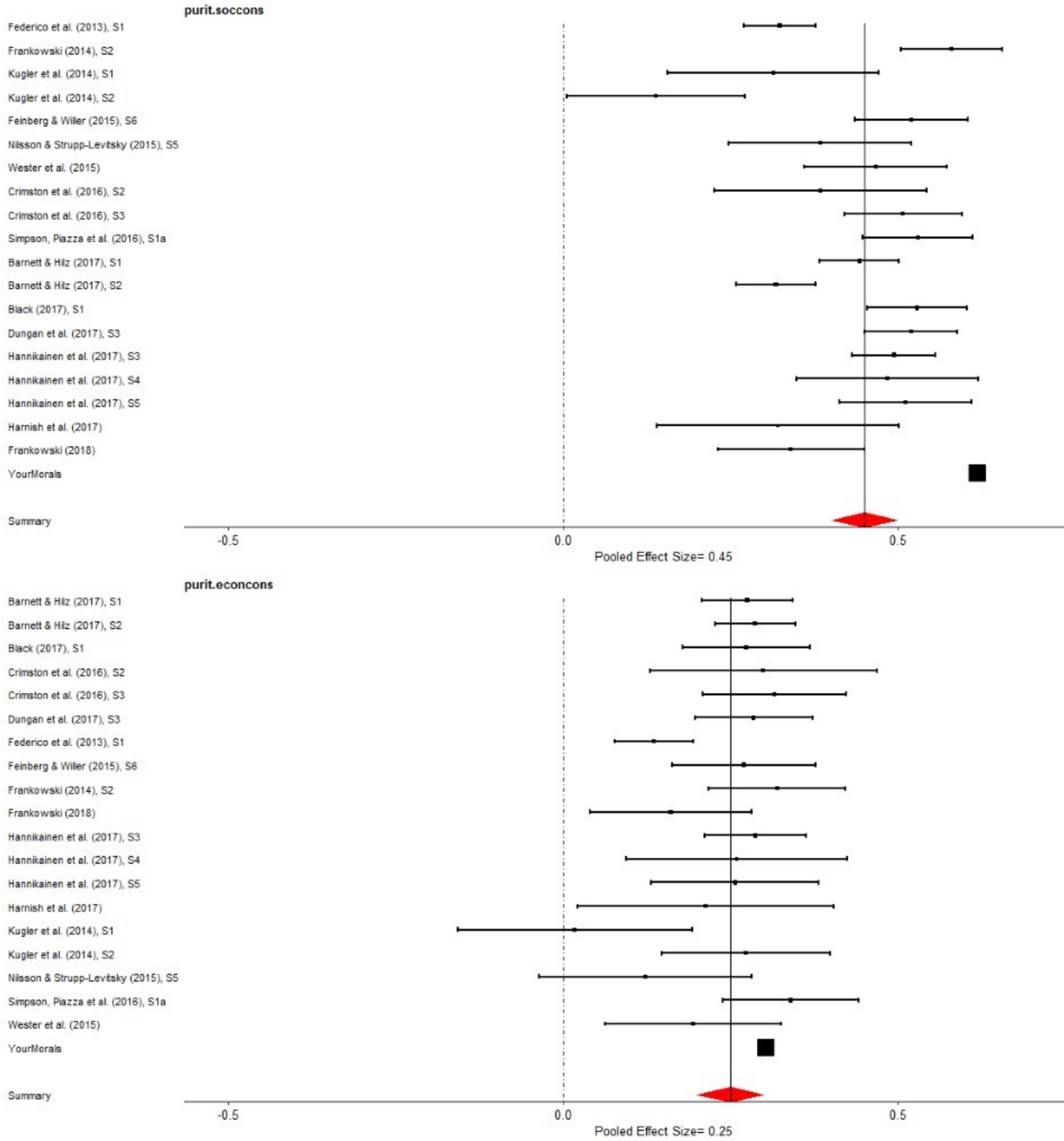

Figure 11. Forest plots for Sanctity-Social conservatism and Sanctity-Economic conservatism.

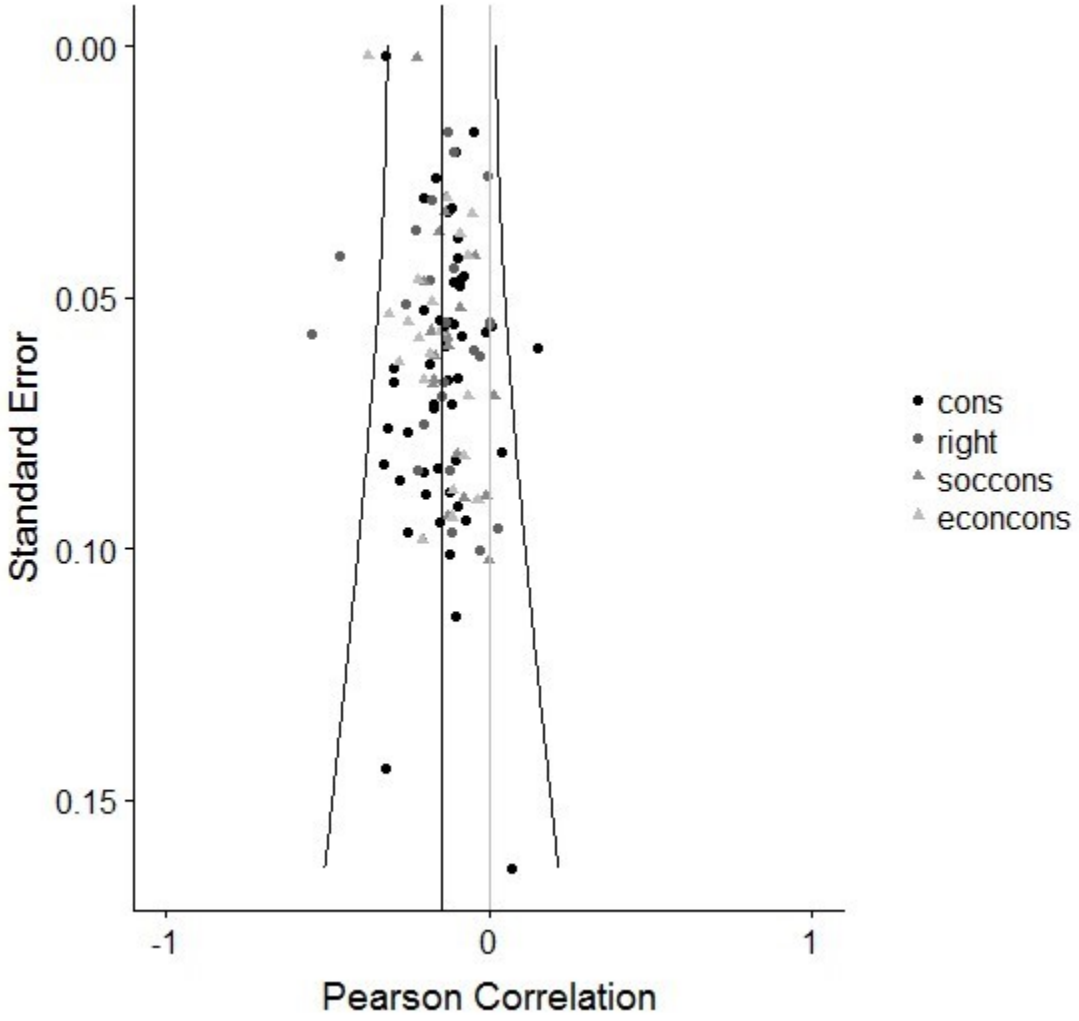

Figure 12. Funnel plot for Care-PSP association.