AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Personality and Social Psychology:
Personality Processes and Individual Differences

ISSN: 0022-3514

2023, Vol. 125, No. 5, 1157–1188
https://doi.org/10.1037/pspp0000470

Morality Beyond the WEIRD: How the Nomological Network of Morality Varies Across Cultures

Mohammad Atari[1], Jonathan Haidt[2], Jesse Graham[3], Sena Koleva[4], Sean T. Stevens[5], and Morteza Dehghani[1]

[1] Department of Psychology, University of Southern California
[2] Stern School of Business, New York University
[3] Department of Management, David Eccles School of Business, University of Utah
[4] User Research, Academia.edu, San Francisco, California, United States
[5] Research and Special Projects Department, The Foundation for Individual Rights in Education, Philadelphia, Pennsylvania, United States

Moral foundations theory has been a generative framework in moral psychology in the last 2 decades. Here, we revisit the theory and develop a new measurement tool, the Moral Foundations Questionnaire–2 (MFQ-2), based on data from 25 populations. We demonstrate empirically that equality and proportionality are distinct moral foundations while retaining the other four existing foundations of care, loyalty, authority, and purity. Three studies were conducted to develop the MFQ-2 and to examine how the nomological network of moral foundations varies across 25 populations. Study 1 ($N = 3,360$, five populations) specified a refined top-down approach for measurement of moral foundations. Study 2 ($N = 3,902$, 19 populations) used a variety of methods (e.g., factor analysis, exploratory structural equations model, network psychometrics, alignment measurement equivalence) to provide evidence that the MFQ-2 fares well in terms of reliability and validity across cultural contexts. We also examined population-level, religious, ideological, and gender differences using the new measure. Study 3 ($N = 1,410$, three populations) provided evidence for convergent validity of the MFQ-2 scores, expanded the nomological network of the six moral foundations, and demonstrated the improved predictive power of the measure compared with the original MFQ. Importantly, our results showed how the nomological network of moral foundations varied across cultural contexts: consistent with a pluralistic view of morality, different foundations were influential in the network of moral foundations depending on cultural context. These studies sharpen the theoretical and methodological resolution of moral foundations theory and provide the field of moral psychology a more accurate instrument for investigating the many ways that moral conflicts and divisions are shaping the modern world.

Keywords: morality, moral foundations theory, culture, values, ideology

Supplemental materials: https://doi.org/10.1037/pspp0000470.supp

Moral foundations theory (MFT; Graham et al., 2013; Haidt & Joseph, 2004) was designed to explain both the variations and ubiquitous aspects of moral judgments across cultures. Specifically, MFT proposed five universally available but contextually variable moral concerns: care/harm, fairness/cheating, loyalty/betrayal, authority/subversion, and purity/degradation.[1] Graham et al. (2009, 2011) developed the Moral Foundations Questionnaire (MFQ) to address the need for a valid and reliable measure of the degree to which people endorse each of these five foundations. This self-report measure has been used in hundreds of empirical

[1] These foundations have come with other names too. Haidt and Graham (2007) referred to them as harm/care, fairness/reciprocity, ingroup/loyalty, authority/respect, and purity/sanctity.

This article was published Online First August 17, 2023.

Mohammad Atari https://orcid.org/0000-0002-4358-7783

The authors thank Joe Henrich, Daphna Oyserman, Mark H. C. Lai, Hajar Yazdiha, Doug Medin, Peter Ditto, Adam B. Cohen, Jonathan Schulz, Tage S. Rai, Morality and Language Group at University of Southern California, and Hot Cognition Lab at University of California, Irvine, for their feedback on earlier versions of this article. This research was partly funded by the National Science Foundation CAREER BCS-1846531 to Morteza Dehghani.

All data, materials, and code are publicly available at https://osf.io/srtxn/.

Mohammad Atari played a lead role in conceptualization, data curation, formal analysis, investigation, methodology, validation, visualization, writing–original draft, and writing–review and editing. Jonathan Haidt played a lead role in conceptualization and supervision, a supporting role in writing–review and editing, and an equal role in investigation and resources. Jesse Graham played a supporting role in conceptualization and an equal role in investigation, supervision, and writing–review and editing. Sena Koleva played a supporting role in writing–review and editing and an equal role in supervision. Sean T. Stevens played a supporting role in conceptualization, data curation, and writing–review and editing. Morteza Dehghani played a lead role in funding acquisition, project administration, resources, and supervision; a supporting role in writing–review and editing; and an equal role in conceptualization, data curation, and validation.

Correspondence concerning this article should be addressed to Mohammad Atari, Department of Human Evolutionary Biology, Harvard University, 11 Divinity Avenue, Cambridge, MA 02138, United States. Email: matari@fas.harvard.edu

studies in different social and behavioral fields, and across dozens of cultures.

However, recent theoretical critiques of MFT and psychometric examinations of the MFQ in diverse samples call for theoretical refinement and psychometric improvement of the questionnaire. To answer these calls, we describe the development of the Moral Foundations Questionnaire–2 (MFQ-2), based on an updated theoretical view on the number of foundations and their content. We develop MFQ-2 using a new item pool administered across 25 populations in their local languages. We present the structural validation of the MFQ-2, its relation to political ideology and religiosity, group differences in the endorsement of moral foundations, as well as an examination of cross-cultural similarities and differences. Our theoretical refinement, this new measurement tool, and our analytic approach allow us to show for the first time the wide variance in nomological networks of morality across populations, opening up several avenues of cross-cultural research.

MFT and the MFQ

How can a moral psychological theory account for the content and structure of morality across cultures when people disagree so much, and so viciously, on moral issues even within the same group? Haidt and Joseph (2004) reviewed evolutionary psychology, cultural psychology, and anthropology and proposed five top candidates for being the psychological "foundations" upon which cultures construct their moralities. Since MFT was first described by Haidt and Joseph (2004), its founders have tried to identify "candidate foundations" for which the empirical evidence was strongest. Graham et al. (2013) proposed five criteria for "foundationhood": (a) being common in third-party normative judgments, (b) automatic affective evaluations, (c) cultural ubiquity though not necessarily universality, (d) evidence of innate preparedness, and (e) a robust preexisting evolutionary model.

The first conceptualization of MFT had five foundations, so the bulk of existing empirical work has investigated these foundations (i.e., care, fairness, loyalty, authority, and purity). Importantly, however, the founders of MFT have continuously emphasized "moral pluralism" (i.e., that morality is more than one thing) rather than insisting on a fixed number of foundations. The first two "individualizing" foundations—care and fairness—center around protection of individuals from harm and unfair treatment, whereas the "binding" foundations of loyalty, authority, and purity are focused on preservation of group cohesion, maintaining boundaries for self and group, and binding individuals into larger groups and institutions. These five foundations were consistent with, and expanded upon, several taxonomies of moral concerns, including Fiske's (1992) relational models theory, and Shweder et al. (1997) account of the "three ethics" of autonomy, community, and divinity.

As the first theory-driven measure of MFT, the MFQ (Graham et al., 2011; hereinafter referred to as MFQ-1 for clarity) measures the degree to which individuals endorse, or value, each of the five areas of morality described by MFT (for an overview of other measures of moral foundations, see Graham et al., 2013). Graham et al. (2011) used both exploratory and confirmatory factor analyses to develop the 30-item MFQ-1 which has two parts, each with a different question format: in the "relevance" section, participants are explicitly asked to evaluate how "relevant to your thinking" various concerns are when they "decide whether something is right or wrong" (e.g., "Whether or not some people were treated differently from others" for fairness). In the second "judgments" section, participants are asked how strongly they agree or disagree with specific moral–judgment statements (e.g., "People should not do things that are disgusting, even if no one is harmed" for purity).

Prior work has provided some evidence for the convergent and construct validity of the MFQ-1. For example, Graham et al. (2011) documented that care scores were positively correlated with empathy, generosity, and pacifism; fairness scores were positively associated with valuing social justice and negatively correlated with social dominance; loyalty scores were positively correlated with concerns over national security; authority scores were positively correlated with respect for tradition and right-wing authoritarianism; and finally, purity scores were positively correlated with valuing self-discipline, religious attendance, disgust sensitivity, and unfavorable attitudes toward casual sexual encounters. Scores on the MFQ-1 have also been correlated with political ideology (Graham et al., 2009; Kivikangas et al., 2021; also see Hatemi et al., 2019), emotional reactions to various moral transgressions (Atari, Mostafazadeh Davani, & Dehghani, 2020), religiosity (Yi & Tsang, 2020), vaccine hesitancy (Amin et al., 2017; Reimer et al., 2022), patterns of language use (Kennedy et al., 2021), public policy attitudes (Christie et al., 2019; Clifford & Jerit, 2013; Koleva et al., 2012), voting (Enke, 2020), and charitable giving (Nilsson et al., 2020).

The MFQ-1 has been used in a wide variety of settings to examine group differences and cultural practices. Haidt and Graham (2007) first applied the theory to understand the "culture wars" between political liberals and conservatives in the United States. Liberals (progressives) in the United States have been found to score slightly higher than their conservative counterparts on care and fairness. On the other hand, conservatives tend to score higher on loyalty, authority, and purity than liberals. This finding has since been replicated multiple times (Kivikangas et al., 2021; Klein et al., 2018).

In a similar vein, Koleva et al. (2012) showed that MFQ-1 scores predicted stances on specific politicized "culture war" issues over and above political ideology, age, gender, religious attendance, and interest in politics.

Cultural Roots of Moral Foundations

Nearly 3 decades ago, Shweder and Haidt (1993) called for culturally informed theories of moral cognition. Scholars have argued that moral appraisals differ substantially across individuals, countries,[2] and historical periods. For example, Shweder et al. (1987) showed that in India, among Brahmins, it is thought to be immoral for a son to eat meat or cut his hair during the 10 days that follow the death of his father, because they were violations of purity practices related to "death pollution." However, these practices are perfectly acceptable in Western cultures and they do not represent any disrespect for one's father's memory. MFT makes it easy to link complex and culturally variable practices to a small set of universally available intuitions, thus facilitating a culturally pluralistic approach to moral judgment. Graham et al. (2011) did not provide a

[2] The terms "countries" and "cultures" have been traditionally conflated, but as our cumulative understanding of culture and cultural evolution has become increasingly sophisticated, it is time to move across from this simplification to better understand cultural variations within and across populations.

comprehensive picture of cultural variation in moral foundations, but they did compare participants from Eastern cultures (South Asia, East Asia, and Southeast Asia) with those from Western cultures (United States, United Kingdom, Canada, and Western Europe). Eastern participants showed stronger concerns about loyalty and purity compared with their Western counterparts, and they were only very slightly more concerned about care, fairness, and authority. According to this research, larger cultural differences in loyalty and purity made sense in light of established cultural differences in collectivism (Triandis, 1995) and the role of spiritual–physical purity concerns in everyday life and religious practices, especially in South Asia (Shweder et al., 1997).

The small effect sizes for all the East–West differences suggest that group differences within cultures (e.g., by gender or political ideology) could exceed the observed East–West differences. However, the analyses in Graham et al. (2011), all involved participants recruited on the Western academic website yourmorals.org answering in English, and so these participants were likely not representative of these world areas. Moreover, Iurino and Saucier (2020) examined the measurement invariance of the short form of the MFQ-1 across 27 countries and concluded that the 20-item version of the MFQ-1 did not have the cross-cultural measurement invariance necessary to meaningfully make such comparisons.

Only recently has it become common and easy to examine morality beyond typical Western, educated, industrialized, rich, and democratic (WEIRD; Henrich et al., 2010) samples in psychological research. For example, Atari, Graham, and Dehghani (2020) evaluated the MFQ-1 in Iranian culture, an understudied less-WEIRD context, and followed up by building a bottom-up model of moral values. These authors also compared moral foundations between Iran and the United States, finding that Iranians' raw scores on the MFQ-1 cannot be directly compared with their American counterparts as the two populations differ in the pattern of responding to questionnaire items, again pointing to the lack of invariance in the MFQ-1 scores.

Notably, in the last few years, a plethora of tools and opportunities have become available for cultural psychologists, which were not available when the MFQ-1 was developed a decade ago. Most notably, Muthukrishna et al. (2020) developed and validated a tool and a quantitative method for measuring the psychological and cultural distance between societies, hence creating a distance scale with any population as the point of comparison, sometimes referred to as the WEIRDness cultural distance. Hence, it is now possible to examine how distance from WEIRD societies (typically exemplified by the United States) is associated with moral foundations in different populations. In addition, it has gotten much easier to collect stratified and representative samples online across nations (Litman et al., 2017), which is particularly important given the selection biases associated with crowdsourcing websites such as https://YourMorals.org, which was a major source of validation data for the MFQ-1 (Kivikangas et al., 2021).

Gaps in Theory and Measurement

As reviewed above, MFT has been shown to be a highly generative theoretical framework in multiple fields. However, recent empirical findings have highlighted limitations and boundary conditions that need to be addressed. Scholars from multiple disciplines have rightly criticized MFT for having failed to include moral concerns for equality and ignoring systemic inequalities (Janoff-Bulman & Carnes, 2013). In addition, MFT has yet to take into account people's altruistic willingness to address existing societal inequalities—even at the expense of one's own group within the same society (Janoff-Bulman & Carnes, 2013). More recently, Meindl et al. (2019) and Skurka et al. (2020) argued for the inclusion of proportionality as a potential foundation, since proportionality is conceptually distinct and empirically distinguishable from the original five foundations in the MFQ-1, including fairness. Equity theory (e.g., Adams, 1963, 1965; Homans, 1961) is probably the most well-known theory coming out of this school of thought, arguing that rewards and punishments should be distributed in accordance with recipients' inputs or contribution. Adams's (1965) work led him to conclude that "when [a person] finds that his outcomes and inputs are not in balance in relation to those of others, feelings of inequity result" (p. 280) and that "there can be little doubt that inequity results in dissatisfaction, in an unpleasant emotional state, be it anger or guilt" (p. 283).

Meindl et al. (2019) conducted a psychometric analysis on a diverse set of justice motives and resource redistribution preferences, demonstrating the existence of two separable types of distributive justice beliefs: equality (influenced by a focus on societal well-being) and proportionality[3] (influenced by a focus on societal power). Accordingly, we make the case, based on prior theorization and cumulative empirical work, that MFT (and moral psychology, more broadly) benefits from breaking fairness into equality and proportionality (Rai & Fiske, 2011). We note that an individual's low scores on proportionality would not necessarily mean they are concerned with equality; the two constructs are not different ends of a single spectrum. Practically, people can take merit into account in their decision making while actively caring about reducing inequality in the society (as seen in some economically conservative, socially liberal individuals in the United States).

One of the novel aspects of the present work is theoretically distinguishing between equality and proportionality. We make this theoretically justified distinction to better measure these distinct routes to justice motivation and fairness, generating novel testable ideas, some of which we test in the present article. Rai and Fiske (2011) argued that equality may be understood in terms of enforcing even balance and in-kind reciprocity in social relations, and that it requires equal treatment, equal say, equal opportunity, equal chance, and identical shares. Proportionality, on the other hand, is directed toward ensuring that, in social relations, for each party rewards or punishments are proportional to their costs, contributions, effort, merit, or guilt (Rai & Fiske, 2011). This theoretical refinement of MFT's fairness foundation is intended as a corrective to the often-tacit assumption in moral psychology that fairness boils down to one single conceptualization of (re)distributing resources in the context of social living (Rai, 2018). In short, we argue that fairness beliefs are diverse and heavily contingent on the socioecological contexts and political systems in which people are chronically embedded.

[3] Our use of the term "proportionality" is interchangeable with "equity" consistent with prior work (e.g., Deutsch, 1975; Haidt & Joseph, 2011; Rai & Fiske, 2011). We do not use the term "equity" because it has recently changed its semantic connotation to mean "equality" in population-level literatures and social justice movements where "equity" means equality of outcomes across groups. We use the term "proportionality" throughout this article to avoid confusion.

Our theoretically justified differentiation between equality and proportionality, as well as developing valid measures for both, opens the door to an array of interesting questions within the framework of MFT. For example, this refinement raises questions of what ideological, economic, ecological, cultural, or even historical factors give rise to equality and proportionality. The relationship between these two constructs can be in itself an interesting question as well. For example, based on recent work in formal computational modeling of ecological niche, it can be the case that availability of diverse socioecological niches to individuals within societies (i.e., more complex societies) can cause equality and proportionality to be more "orthogonal" (i.e., more distinguishable constructs rather than one being a special case of the other, or reflecting different aspects of a more basic psychological construct; Durkee et al., 2020; Smaldino et al., 2019). In addition, the strength of social norms and lower individualism may account for higher covariance between equality and proportionality (for a review of trait covariance structures as a function of cultural factors and socioecological environments, see Gurven, 2018).

In terms of measurement of moral foundations, Graham et al. (2011) conducted confirmatory factor analyses (CFAs) based on the English version of the MFQ-1 to determine whether the five-factor model of MFT fits data better than alternative models and demonstrated that the five-factor model fits the data better than the two-factor (individualizing vs. binding foundations) and single-factor models. Furthermore, independent scale-validation studies in different cultural contexts have replicated this initial finding (e.g., Davies et al., 2014; D. E. Davis et al., 2017; Nejat & Hatami, 2019; Nilsson & Erlandsson, 2015; Yalçındağ et al., 2019; Yilmaz et al., 2016). However, in all these studies, fit indices of the five-factor model were substantially below the conventional thresholds. A recent cross-cultural study using the 20-item version of MFQ-1 showed measurement noninvariance across 27 societies (Iurino & Saucier, 2020). In other words, there is some evidence suggesting that the five-factor model proposed by the theory is not valid across these countries, and subscale scores may not be meaningfully compared across populations because patterns of responding are different from one population to another. Internal consistency of MFQ-1 subscales (or foundations) also fails to reach conventional thresholds of 0.70, especially in more diverse or representative samples. In most less-WEIRD populations, the individualizing foundations (care and fairness) fare especially poorly, and loyalty and authority scores typically fail to achieve adequate internal consistency (Nejat & Hatami, 2019). It is not clear whether these measurement issues reflect a generalizability problem on the theoretical side, or a psychometric problem with cross-cultural validity of MFQ-1 as it was not constructed with input from multiple cultures.

Recently, Doğruyol et al. (2019) showed that the five-factor model of moral foundations, operationalized by the short version of the MFQ-1 (20 items), is stable and invariant across WEIRD and non-WEIRD societies; however, these authors used the problematic dichotomy of WEIRD versus non-WEIRD rather than treating societies on a continuum of WEIRDness (Muthukrishna et al., 2020). Atari, Graham, and Dehghani (2020) reported noninvariance of MFQ-1 scores between a less-WEIRD society (Iran) and the United States at configural (i.e., the overall factor structure stipulated by the five-factor model did not hold across populations), metric (i.e., item-factor loadings were not equivalent across populations), and scalar models (i.e., the item intercepts were not equivalent across populations). These authors also had some difficulty in translating some items into local languages (e.g., the item "I would call some acts wrong on the grounds that they are unnatural"), arguing that while MFT is a useful theoretical framework in less-WEIRD societies, MFQ-1 scores may not be reliable and predictive of social behaviors. In addition, Atari, Graham, and Dehghani (2020) used network psychometric methods and found that regardless of mean endorsement of moral foundations, the network of items and foundations are substantially different between the two countries, with Iran having a denser interconnected network of moral foundations, compared with the more segregated network of moral concerns in the United States, wherein care–fairness and loyalty–authority–purity are two disconnected "islands" (or subnetworks). Another study in the United Kingdom also failed to replicate the five-factor model originally proposed by Graham et al. (2011) and suggested that "compassion" and "traditionalism" may account for the structure of the MFQ-1 in the United Kingdom (Harper & Rhodes, 2021). A recent factor-analytic study with young Muslim adults in Pakistan also found that the MFQ-1 has psychometric limitations and cannot be reliably used to measure the original five moral foundations in Punjab (Akhtar et al., 2023).

Overview of the Present Research

In the decade since the development of its gold standard measure, MFT has substantially broadened the range of moral concerns studied in moral psychology by encouraging researchers to look beyond harm and fairness (Graham et al., 2018). But substantial investigations in cultural differences in moral priorities require revision of both the theory and the measure. In this work, we have five major goals. First, we refine MFTs view on fairness by introducing equality and proportionality as novel and distinct foundations. Second, we generate a completely new item pool and develop the MFQ-2 across populations using local languages and generalizable samples. Third, we examine the structural validity and measurement invariance of the MFQ-2 across cultures. Fourth, we examine group differences (population-level, ideological, gender, and religious differences) using the novel MFQ-2, conceptually replicating prior work that has established these differences. Fifth, we establish external validity of the MFQ-2 by examining associations with other scales meant to capture similar constructs.

Our measurement philosophy follows recommendations by Flake et al. (2017) in following three phases of measure development: substantive (Phase 1: literature review, construct conceptualization, item pool development); structural (Phase 2: item analysis, factor analysis, reliability, measurement invariance); and external (Phase 3: convergent validity, group differences). Our five studies come in three phases, which we summarize in Table 1.

Study 1a

Study 1a was conducted to define the top-down structure we intend for MFQ-2 (care, equality, proportionality, loyalty, authority, and purity) and to develop a preliminary MFQ-2 item pool that could be used to operationalize this theory-driven model. Ideally, this item pool should be broad and balanced, with each foundation represented by several candidate items. Our conceptual definitions of the six foundations we aim to measure are shown in Table 2. In all studies,

Table 1
Description of Studies, Aims, and Samples

Phase	Study	Description	Sample size	Nations
1	1a	Literature review, panel discussion, item pool development	840	2
1	1b	Panel discussion, item pool reduction	971	3
1	1c	Panel discussion, item pool reduction	1,549	3
2	2	Factor analysis, reliability, measurement invariance, group differences	3,902	19
3	3	Convergent validity	1,410	3

we have data from at least two nations in order to avoid focusing narrowly on one particular "default" culture. To avoid the "home-field disadvantage" (Medin et al., 2010), we also made sure that our team has a diverse set of cultural backgrounds and views to make sure that our item pool was not Eurocentric or biased toward a particular ideology. Here, we describe the process of generating the item pool, initial analyses, and reducing the item pool for use in the next studies.

Method

Participants and Procedure

We aimed to recruit 1,000 participants from the United States and India using Cloud Research (Litman & Robinson, 2020). After removing participants who failed any of our four attention checks, 840 participants remained for statistical analyses (India: $n = 346$; United States: $n = 494$). The distribution of participants based on their Internet Protocol address is shown in Figure 1. All participants first completed the item pool (see Measures section), then they completed the MFQ-1, and finally reported their demographic details. The present sample ranged in age from 18 to 77 years old ($M = 34.24$, $SD = 11.02$), and included an approximately equal number of men and women (55.83% male). Most of our U.S. sample identified as White (71.3%), followed by Asian (16.0%), Hispanic or Latino/Latinx (11.3%), and African American (9.3%). Our sampling strategy and exploratory analysis plan were preregistered on the Open Science Framework (OSF; https://osf.io/3hefa). This study was conducted in August 2020. The Institutional Review Board (IRB) at the University of Southern California approved this set of studies (UP-20-00570).

Table 2
Conceptual Definitions of Six Moral Foundations

Foundation	Definition
Care	Intuitions about avoiding emotional and physical damage to another individual.
Equality	Intuitions about equal treatment and equal outcome for individuals.
Proportionality	Intuitions about individuals getting rewarded in proportion to their merit or contribution.
Loyalty	Intuitions about cooperating with ingroups and competing with outgroups.
Authority	Intuitions about deference toward legitimate authorities and the defense of traditions, all of which are seen as providing stability and fending off chaos.
Purity	Intuitions about avoiding bodily and spiritual contamination and degradation.

Measures

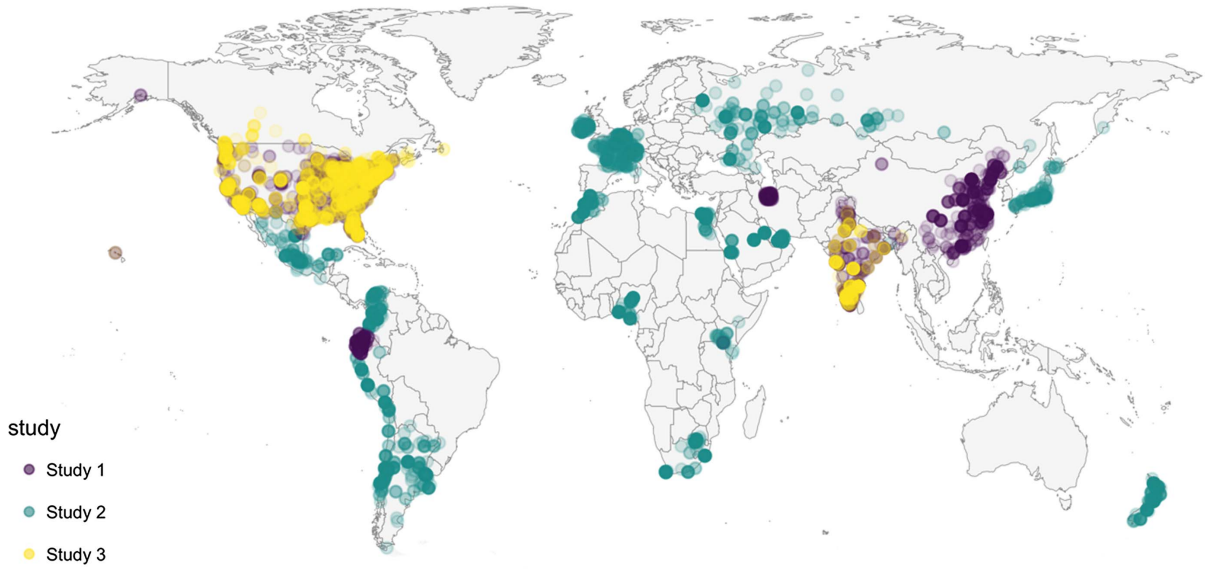
The measures used in the present study are described below.

MoFQ-2 Item Pool. We reviewed the extant MFT literature, as well as criticisms regarding some of the items in MFQ-1 (see Table 1). We aimed to develop an initial item pool with over 100 items all in a declarative form, similar to the "judgments" part of the MFQ-1. Since the "relevance" items have been shown to reduce internal consistencies and cause confusion among some researchers (e.g., by only using relevance items rather than using both relevance and judgments, a practice that should be avoided), we decided to drop the relevance format. The first author initially prepared 20–25 items per foundation, and then other authors added more items to the preliminary pool. All authors met 7 times to finalize an item pool of 116 items. While there was some disagreement regarding some items, all authors agreed that these 116 items are acceptable candidates to measure six foundations (care: 15 items; equality: 18 items; proportionality: 25 items; loyalty: 19 items; authority: 20 items; and purity: 19 items). Since MFT is based on the social intuitionist model (Haidt, 2001) indicating that moral evaluations occur rapidly and automatically, arising from effortless and heuristic processing (known as System 1 thinking in psychology; Kahneman, 2011), we intentionally included a large number of items that include emotional reactions, self-perceived emotional tendencies, and emotional displays.

While we no longer have a "fairness" subscale in MFQ-2, some items did not clearly belong to either equality or proportionality in the first round of data collection (e.g., "When the government makes laws, the number one principle should be ensuring that everyone is treated fairly"). Therefore, we left these items as they were to explore how they relate to new equality and proportionality items. The response option was provided from 1 (*does not describe me at all*) to 5 (*describes me extremely well*) based on our qualitative examination of different response options and consultation with survey researchers, who noted problems with the previous agree–disagree response options (see Krosnick & Fabrigar, 1997). In this study, we also provided the option for all participants to give feedback if any of the items were not comprehensible, did not read well, or were otherwise unclear. We also made sure that items included both negative and positive poles (or virtues and vices) for each foundation (e.g., punishing cheaters vs. rewarding merit, for proportionality).

MFQ. All participants completed the 30-item MFQ-1 (Graham et al., 2011), which consists of two 15-item parts, relevance and judgments. The relevance part measures the five foundations using the relevance individuals ascribe to each of the foundations. Items in the relevance section are rated along a 6-point Likert-type scale ranging from 0 (*not at all relevant*) to 5 (*extremely relevant*). The judgments section has contextualized items that can gauge moral

Figure 1
The Geographical Distribution of Participants in Three Studies



Note. We did not have Internet Protocol information from our Iranian sample, so we are representing all participants' geolocation on Tehran in this map. See the online article for the color version of this figure.

judgments related to the five moral foundations. Items in the judgments section are rated along a 6-point Likert-type scale ranging from 0 (*strongly disagree*) to 5 (*strongly agree*). The internal consistency coefficients were .66, .64, .79, .78, and .86 for care, fairness, loyalty, authority, and purity, respectively.

Political Ideology. All participants rated their political affiliation with the Republican Party or the Democratic Party along a 7-point scale ranging from 1 (*strong Democrat*) to 7 (*strong Republican*). Participants were also asked to rate their political conservatism on a scale ranging from 1 (*very liberal*) to 7 (*very conservative*). We then averaged these two ratings to create a political-orientation score, wherein higher scores indicated a more conservative political orientation. A similar method was used in previous work for assessment of political ideology (Jost & Thompson, 2000). The internal consistency of these two items was high in the American sample ($\alpha = .90$). In the Indian subsample, we only use the conservatism item to quantify political ideology.

Religiosity. All participants self-reported their religious affiliation, as well as religiosity using the single item "On a scale from 0 to 10, how religious do you consider yourself?" Participants indicated the level of religiosity using a slider ranging from 0 to 10. The Indian subsample was significantly more religious ($M = 6.73$, $SD = 2.69$) than the American subsample ($M = 3.99$, $SD = 3.59$, $t = 12.66$, Welch-corrected $df = 834.08$, $p < .001$).

Results and Discussion

We first examine descriptive statistics for each item. Some items had floor or ceiling effects, indicated by high skewness. These items were considered for discarding. We also performed a number of different factor analyses and reliability analyses to see which items hold together well while keeping the breadth of each foundation. Specifically, we conducted exploratory factor analyses (EFAs) for items belonging to each foundation. For care, we dropped three items as they did not reach the adequate 0.4 item-factor loading criterion, hence 12 items were selected to be used in Study 1b. Based on similar criteria about item-factor loadings, qualitative analysis of items, and feedback from participants and all authors, 14, 18, 16, 15, and 15 items were selected to be administered in Study 1b for equality, proportionality, loyalty, authority, and purity, respectively. Cross-societal differences in each item, as well as correlations between these items and relevant MFQ-1 subscales, are presented in Supplemental Materials. Therefore, we reduced our initial, crude item pool of 116 items to a sharper and more focused set of 90 items for further data collection and analysis in Study 1b.

Study 1b

Study 1b was conducted to refine the 90-item pool from Study 1a into a more fine-grained MFQ-2 pool. To do this, we administered the 90 items to a diverse sample of participants from three nations, namely, India, the United States, and Iran. We specifically chose Iran because MFQ-1's structure was particularly inconsistent with the structure typically observed in Western cultures (Atari, Graham, & Dehghani, 2020) and because Iran is approximately culturally equidistant from both WEIRD populations (e.g., the United States) and developed Eastern Asian populations (e.g., China; Muthukrishna et al., 2020). We used these data to select the next set of MFQ-2 items and conduct a preliminary examination of the MFQ-2's psychometric properties.

Method

Participants and Procedure

We aimed to recruit 1,000 participants from the United States and India using Cloud Research (Litman & Robinson, 2020). We

translated all 90 items into Farsi using the standard back translation procedure (Brislin, 1970) and aimed to recruit Iranian participants by advertising the survey link on social media platforms.[4] After removing participants who failed any of our four attention checks, 971 participants remained for statistical analyses (India: $n = 380$; United States: $n = 491$; Iran: $n = 100$). The current sample ranged in age from 18 to 77 years old ($M = 34.81$, $SD = 16.92$), and included an approximately equal number of men and women (53.26% male). Most Americans identified as White (69.2%), followed by African American (12.8%), Asian (11.8%), and Hispanic or Latino/Latinx (10.0%; participants could select multiple race/ethnicities). Our sampling strategy and exploratory analysis plan were preregistered on OSF (https://osf.io/d2b6s). This study was conducted in September 2020.

Measures

All participants completed the 90-item pool of the MFQ-2 finalized in Study 1a (care: 12 items; equality: 14 items; proportionality: 18 items; loyalty: 16 items; authority: 15 items; purity: 15 items). They then self-reported their political ideology, religiosity, and demographic details. For political ideology, we used the two-item measure in Study 1a ($\alpha = .88$) and used the single-item measure of liberalism–conservatism in India and Iran (with slight wording modification in Farsi for cultural fluency).

Results and Discussion

As in Study 1a, we examined all items' descriptive statistics, checking potential ceiling or floor effects in any of the cultures we had data from. After item analysis, we conducted foundation-level EFAs across populations (see Supplemental Materials). After item analysis and EFAs, 19 items were discarded overall, leaving 71 items for administration in Study 1c. Cross-societal differences in each item and factor analyses are presented in Supplemental Materials. In this study, we reduced our item pool to 71 items for further analysis in Study 1c while making sure that items hold together well (while maintaining breadth of conceptual coverage) and correlate with relevant MFQ-1 foundations reasonably.

Study 1c

In Study 1c, we prepare the final item pool for our main cross-cultural data collection effort in Study 2. We administer the 71-item pool in three populations in order to further reduce the number of items. Here, we chose three populations with the highest feasible cultural distance in order to maximize the diversity of our samples. Based on Muthukrishna et al. (2020) WEIRD cultural distance, we chose the United States, Ecuador, and China. Ecuador is culturally distant from both the United States and China, is a Spanish-speaking country with relatively high diversity in people's languages and subcultures, and remains one of the most understudied cultures in moral psychology. Since the geography of Ecuador is very diverse, its population's lifestyles, principal work, and economic structure are also diverse. There are fishermen along the coasts, cattlemen in the southern highlands, farmers on central highland slopes, and oil workers in the Amazon (Cruza-Guet et al., 2009). In addition, here we address one of the important limitations of our samples in Studies 1a and 1b, that is, relying on convenience sampling. Here, we recruit stratified national samples mirroring national demographics in terms of gender, education, and age (and political ideology in the United States). In addition, in this study, we used psychometric network methods to diversify our methodological toolbox while choosing the best-performing items (Christensen et al., 2020).

Method

Participants and Procedure

We aimed to recruit about 1,500 participants from the United States, Ecuador, and China using Qualtrics Panels. We translated all items into Spanish and Mandarin using the back translation technique (Brislin, 1970). Two independent bilingual researchers double-checked the final items for cultural fluency. Participants who failed any of the three attention checks were immediately dropped and replaced by Qualtrics Panels, in order to achieve stratified samples in terms of age, gender, race, and political orientation. Overall, 1,549 participants remained for statistical analyses (United States: $n = 515$; Ecuador: $n = 517$; China: $n = 517$). All participants first completed the 71-item pool, then they completed MFQ-1, and finally reported their demographic details. Participants also completed a few items at the end of the survey, related to another project. The current sample ranged in age from 18 to 87 years old ($M = 40.92$, $SD = 16.02$), and included an approximately equal number of men and women (49.9% male). Most Americans identified as White (73.8%), followed by African American (8.7%) and Asian (3.9%). Our exploratory analysis plan was preregistered after data collection, but before any data analysis, available on OSF (https://osf.io/qae9c). This study was conducted in January 2021.

Measures

All participants completed the 71-item pool of MFQ-2 (care: 10 items; equality: 10 items; proportionality: 13 items; loyalty: 13 items; authority: 12 items; purity: 13 items). As in Studies 1a and 1b, they then completed MFQ-1 (α coefficients ranged between .62 [fairness] and .78 [purity]), political ideology, religiosity, and demographic details. For political ideology, we used the two-item measure in Studies 1a and 1b ($\alpha = .72$), and used a single-item measure of conservatism in Ecuador and China.

Results and Discussion

As in Studies 1a and 1b, we examined all 71 items' descriptive statistics and conducted foundation-level EFAs (see Supplemental Materials). Since our aim was for MFQ-2 to have six items per foundation (similar to the MFQ-1), we aimed to select seven to nine items for each foundation. Our aim in this study was to combine item analysis, factor analysis, psychometric network analysis, and qualitative examination of the breadth for each foundation's items to avoid redundancy. We wanted the final measure to adequately represent each foundation's considerable bandwidth—rather than narrowing the range of moral concerns assessed—in order to maintain the MFQ-2's descriptive and predictive breadth. For care, we dropped two items based on our qualitative examination of remaining items, leaving eight items for Study 2. For equality, two

[4] We preregistered 200 participants to be recruited from Iran, but since the survey was relatively long and we could not compensate participants due to economic sanctions, we stopped data collection at 100 participants.

items were discarded for having an item-factor loading problem in at least one nation, and one item was discarded for similarity to another item, leaving seven items for Study 2. For proportionality, two items were discarded based on psychometric network analysis and three items were dropped to increase item diversity, leaving eight items for inclusion in Study 2. For loyalty, we dropped three items based on EFA results, and discarded two items to reduce content redundancy. In addition, we added one new item to loyalty items ("It is more important to be a good team member than to express oneself") to test whether it can hang with other items in a desirable way, leaving a total of nine items for inclusion in Study 2. For authority, we discarded one item based on EFA results and discarded two items due to content redundancy with other existing items, leaving nine candidate items for Study 2. Finally, for purity, we discarded three items based on qualitative examination of items' content and discarded one item due to centrality issues in the psychometric network analysis, leaving nine items for administering in Study 2. Hence, in this study, we prepared 50 candidate items for translation and use in Study 2 across populations, aiming for the final MFQ-2 to have 36 balanced items. Cross-societal differences in the foundations based on these items, as well as correlations between these foundations, MFQ subscales, and political ideology, are presented in Supplemental Materials.

We believe that the psychometric network approach to moral foundations (used here and in Study 2) has great promise for the study of moral foundations. First, in the network approach, individual moral foundations are not obscured by the whole network of moral foundations. Hence, this approach offers a more holistic representation of the moral domain through intuitive and insightful visualization of the links between the foundations (Bringmann & Eronen, 2018). Second, the network approach enables us to examine "centrality" (see Costantini et al., 2015). Since foundations are allowed to associate with one another in a whole network, the question of which foundations are more central (vs. peripheral) becomes salient. Central foundations would be meaningfully connected to other foundations associated with the respective moral underpinning. Therefore, the network approach provides the opportunity to identify which foundations (and items) are leading to these interfoundation (and interitem) associations, which guides scale developers in deciding whether or not to include certain items based on the construct they intend to quantify (Christensen et al., 2020). Another kind of question that can be asked using a network approach is whether changes in one foundation would cause a meaningful change in another foundation. This idea has been popular among lay people, philosophers, and psychologists alike, but has remained largely unexamined. Finally, several researchers have recommended using latent modeling approaches (e.g., factor analysis) and network analysis as complementary methods (Burger et al., 2022). For example, Clifton and Webster (2017, p. 451) suggested that "[the network approach] has the potential to integrate and advance both the methods and theories used in social and personality psychology."

Study 2

Our goal in Study 2 was to (a) finalize the MFQ-2 items based on cross-cultural data; (b) establish measurement invariance across groups and examine group differences in MFQ-2 scores; and (c) examine the variation of the nomological network of moral foundations across populations. In addition, we examined which moral foundations are more central in each population. We administered the final 50-item pool from Study 1 to a diverse sample of participants from 19 new populations, none of which were sampled in Study 1.

Method

Participants and Procedure

We aimed to recruit stratified samples in terms of age, gender, and political orientation from diverse cultural backgrounds. Based on Muthukrishna et al. (2020) cultural distance metric, we made a list of candidate nations. We then cross-referenced that with the feasibility of stratified data collection administered by Qualtrics Panels targeting 200 participants per nation. We collected nationally stratified samples from 19 nations ($N = 3{,}902$). Details about these samples are provided in Table 3. All measures were translated into local languages using a third-party professional translating company. Then, independent bilingual researchers checked the translations and made sure of the fluency of all items. Discrepancies and modifications were addressed between the translation companies, independent researchers, and the authors. All participants completed the 50-item pool and a few demographic questions. Participants who failed any of the three attention checks were terminated from continuing the survey. This study was conducted in May 2021. Our university's IRB approved the study (UP-20-00570).

Measures

All participants first completed a few demographic questions: country of residence, age, gender, and political ideology. Then, they completed the 50-item pool of MFQ-2 prepared in Study 1. The order of questions was randomized. Participants then completed some measures unrelated to this study, a single-item measure of religiosity, and demographic details. For political ideology, we used a single-item measure, rated along a 10-point scale, that can work equally well across nations ("In political matters, people talk of 'the left' and 'the right.' How would you place your views on this scale, generally speaking?"). A few other items, related to another project, were also included at the end of the survey.

Analytic Strategy

Our statistical analyses of the data come in three separate but related parts. In Part 1, we use the exploratory structural equations modeling (ESEM) framework (Asparouhov & Muthén, 2009; Marsh et al., 2014), as well as descriptive item analysis in order to finalize the 36-item MFQ-2. ESEM is a synergy of EFA and CFA, incorporating the advantages of both EFA and CFA. ESEM is an effective method in the psychometric examination of multidimensional measures and can easily be complemented with other modeling techniques. In the presence of multidimensionality resulting from the assessment of conceptually related constructs (Morin et al., 2016), it is possible that the restrictive assumptions of CFA are violated, and ESEM models may outperform CFAs. In our EFA, we used an EFA with six dimensions, using the maximum likelihood factoring method and oblique rotation. In the second part, we conduct measurement invariance across all populations. To test measurement invariance, we use the multigroup factor analysis

Table 3
Description of Samples Across 19 Nations in Study 2

Nation	*n*	% female	Age *M* (*SD*)	Language	Sample's majority religion (%)	WEIRD cultural distance [95% CI]
Argentina	205	48.8%	42.5 (15.0)	Spanish	Christianity (62.0%)	.071 [.069, .075]
Belgium	205	49.8%	45.1 (17.0)	French	Christianity (47.8%)	NA
Chile	205	49.3%	42.4 (16.2)	Spanish	Christianity (58.5%)	.078 [.075, .081]
Colombia	205	48.8%	41.0 (15.0)	Spanish	Christianity (64.4%)	.102 [.099, .106]
Egypt	205	49.8%	44.8 (16.8)	Arabic	Islam (94.1%)	.234 [.228, .241]
France	206	49.0%	43.7 (16.9)	French	Christianity (48.5%)	.079 [.075, .085]
Ireland	205	50.2%	44.8 (16.7)	English	Christianity (66.3%)	NA
Japan	207	49.3%	47.2 (15.3)	Japanese	None (46.9%)	.115 [.112, .119]
Kenya	205	48.3%	37.6 (12.4)	English	Christianity (85.4%)	NA
Mexico	206	46.6%	41.9 (15.4)	Spanish	Christianity (53.4%)	.077 [.074, .080]
Morocco	205	48.3%	41.8 (14.7)	Spanish	Islam (96.6%)	.149 [.145, .155]
New Zealand	205	48.3%	47.4 (18.2)	English	None (47.3%)	.053 [.050, .058]
Nigeria	205	41.4%	39.1 (13.6)	English	Christianity (76.6%)	.130 [.126, .135]
Peru	205	37.6%	37.0 (13.8)	Spanish	Christianity (62.9%)	.090 [.087, .094]
Russia	206	45.6%	41.7 (14.9)	Russian	Christianity (62.6%)	.085 [.083, .088]
Saudi Arabia	207	48.3%	42.4 (15.5)	Arabic	Islam (96.1%)	NA
South Africa	205	47.3%	41.3 (15.4)	English	Christianity (81.0%)	.076 [.073, .079]
Switzerland	205	50.2%	46.7 (16.8)	French	Christianity (52.7%)	.068 [.064, .074]
UAE	205	49.3%	43.1 (14.7)	Arabic	Islam (84.9%)	NA

Note. WEIRD cultural distance is based on Muthukrishna et al. (2020) and represents cultural distance from the United States. This index is a robust method of measuring cultural distance, grounded in evolutionary theory. A distance of 0 means that the populations are identical, and if the two equal-size populations are more homogeneous but different, we get the maximum distance of 1 (for details, see Muthukrishna et al., 2020). NA = not available; CI = confidence interval; WEIRD = Western, educated, industrialized, rich, and democratic.

alignment method (or simply, "alignment"), which has been developed as an effective and novel method to test metric and scalar invariance (Asparouhov & Muthén, 2014). This method aims to address issues in multigroup confirmatory factor analysis (MGCFA) invariance testing, such as difficulties in establishing exact scalar invariance with many groups (as is the case in the current work). The key distinction between MGCFA and alignment is that alignment does not require equality restrictions on factor loadings and intercepts across groups. The base assumption of the alignment method is that the number of noninvariant measurement parameters and the extent of measurement noninvariance between groups can be held to an acceptable minimum by producing a solution that features many approximately invariant parameters and few parameters with large noninvariances (Fischer & Karl, 2019). Our ultimate goal is to compare latent factor means of moral foundations across groups (here, populations); therefore, the alignment method estimates factor loadings, factor means, factor variances, and item intercepts (Asparouhov & Muthén, 2014).

Notably, a sample size of 200 per group is adequate to conduct measurement invariance analysis given the number of items per foundation, item commonalities, and estimated factor loadings in Study 1 (see Meade & Bauer, 2007). After measurement invariance is evidenced, we compare and contrast populations across the six dimensions of MFQ-2. We also examine the relationship between MFQ-2 scores and WEIRDness cultural distance scores (Muthukrishna et al., 2020). In the third part, we examine gender, religious, and ideological differences. To do so, we rely on multilevel models wherein participants are modeled as nested within groups.

We then proceed to examine how the psychometric network of the six foundations varies across populations. In psychological networks, there are nodes that represent observed variables and edges that represent statistical associations. We used exploratory graph analysis (EGA; Golino & Epskamp, 2017) to estimate the number of higher order dimensions in MFQ-2 scores. A typical way of assessing the importance of nodes in psychometric networks is to compute *centrality* measures of the network. Centrality may be considered an umbrella term that reflects how well-connected a node is to the rest of the network (Clifton & Webster, 2017). Here, we use measures of centrality, that is, a node's influence in the network using the "strength" index (i.e., how well a node is connected to other nodes), which statistically denotes the sum of the weights connected to each node (Burger et al., 2022; Epskamp et al., 2018).

Results and Discussion

Exploratory Structural Equations Models

We first conducted an ESEM on the entirety of the data (comparative fit index [CFI] = .958, Tucker–Lewis index [TLI] = .958, root-mean-square error of approximation [RMSEA] = .029, standardized root-mean-square residual [SRMR] = .027) and discarded 14 items for having cross-loadings, while making sure that remaining items are not redundant in content. These results are presented in Supplemental Materials. We then conducted a secondary ESEM with the final 36 items on the whole data (CFI = .979, TLI = .978, RMSEA = .024, SRMR = .023). All items and loadings are presented in Table 4. We then conducted the same model using MGCFA and found the model to fit the data well across populations (CFI = .896, TLI = .893, RMSEA = .052, SRMR = .070). Hence, we found strong evidence for configural invariance, that is, the same six-dimensional factorial structure holds across all samples (for country-specific CFAs, see Supplemental Materials). Accordingly, the final 36-item MFQ-2 has good structural validity across populations (we note that some purity items have loadings

Table 4
Results of Exploratory Structural Equations Modeling (Study 2)

Item	Care	Equality	Loyalty	Authority	Purity	Proportionality
It pains me when I see someone ignoring the needs of another human being.	**0.66**	0.05	0.09	0.04	−0.02	−0.04
I am empathetic toward those people who have suffered in their lives.	**0.70**	0.04	0.00	−0.01	0.01	0.05
I believe that compassion for those who are suffering is one of the most crucial virtues.	**0.73**	−0.01	0.00	−0.04	0.08	0.01
Caring for people who have suffered is an important virtue.	**0.73**	−0.01	0.02	−0.01	0.05	0.04
We should all care for people who are in emotional pain.	**0.74**	0.03	0.00	0.00	0.05	−0.01
Everyone should try to comfort people who are going through something hard.	**0.64**	0.05	0.04	0.07	0.03	0.00
I believe it would be ideal if everyone in society wound up with roughly the same amount of money.	0.01	**0.81**	0.00	−0.05	−0.01	0.01
When people work together toward a common goal, they should share the rewards equally, even if some worked harder on it.	0.07	**0.38**	0.06	0.26	−0.02	−0.16
I believe that everyone should be given the same quantity of resources in life.	0.23	**0.54**	−0.02	0.07	−0.03	0.02
The world would be a better place if everyone made the same amount of money.	−0.04	**0.88**	−0.03	0.00	0.01	0.00
I get upset when some people have a lot more money than others in my country.	0.07	**0.52**	0.09	−0.3	0.07	0.12
Our society would have fewer problems if people had the same income.	−0.03	**0.86**	−0.04	0.04	0.00	−0.02
I feel good when I see cheaters get caught and punished.	0.09	0.02	0.20	0.03	0.00	**0.29**
I think people should be rewarded in proportion to what they contribute.	0.08	−0.04	0.03	0.06	0.09	**0.54**
I think people who are more hardworking should end up with more money.	0.00	0.02	0.01	0.01	0.04	**0.72**
It makes me happy when people are recognized on their merits.	0.32	−0.07	−0.07	0.40	−0.09	**0.27**
In a fair society, those who work hard should live with higher standards of living.	−0.06	0.04	0.10	−0.04	0.01	**0.72**
The effort a worker puts into a job ought to be reflected in the size of a raise they receive.	0.13	0.06	−0.04	0.10	−0.04	**0.53**
I think children should be taught to be loyal to their country.	0.00	0.00	**0.78**	0.12	−0.01	−0.01
I believe the strength of a sports team comes from the loyalty of its members to each other.	0.19	0.03	**0.10**	0.36	−0.04	0.13
Everyone should love their own community.	0.15	0.06	**0.37**	0.21	0.08	0.01
Everyone should defend their country, if called upon.	−0.02	−0.01	**0.70**	0.02	0.08	0.07
Everyone should feel proud when a person in their community wins in an international competition.	0.21	−0.03	**0.27**	0.28	−0.03	0.08
It upsets me when people have no loyalty to their country.	0.04	0.00	**0.83**	−0.01	−0.04	−0.02
I feel that most traditions serve a valuable function in keeping society orderly.	−0.03	0.07	0.21	**0.39**	0.10	0.08
I think having a strong leader is good for society.	0.10	−0.09	0.09	**0.32**	0.14	0.14
I think it is important for societies to cherish their traditional values.	−0.04	0.05	0.22	**0.44**	0.06	0.08
I believe that one of the most important values to teach children is to have respect for authority.	−0.02	0.00	0.16	**0.59**	0.05	−0.03
I think obedience to parents is an important virtue.	0.01	0.00	0.09	**0.55**	0.19	0.04
We all need to learn from our elders.	0.09	0.07	0.10	**0.49**	0.05	0.00
I believe chastity is an important virtue.	0.05	−0.02	0.00	−0.01	**0.84**	0.02
I think the human body should be treated like a temple, housing something sacred within.	0.17	0.05	−0.02	0.37	**0.21**	−0.01
I admire people who keep their virginity until marriage.	−0.01	0.01	−0.03	0.09	**0.79**	−0.01
People should try to use natural medicines rather than chemically identical human-made ones.	0.06	0.23	0.06	0.10	**0.25**	0.00
If I found out that an acquaintance had an unusual but harmless sexual fetish I would feel uneasy about them.	−0.04	0.10	0.12	−0.05	**0.48**	0.03
It upsets me when people use foul language like it is nothing.	0.11	−0.01	0.28	0.05	**0.29**	−0.01

Note. Relevant item-factor loadings are in bold. Note than due to the nature of exploratory structural equations modeling, some items may be loaded more strongly on other factors.

smaller than the conventional cutoff value of .30 on the purity factor, and cross-loadings on other factors).

Reliability of MFQ-2

Various reliability estimates have been proposed in the literature, with the coefficient alpha (α) being the most prominent. However, coefficient α ignores the measure's internal factor structure, which should be inherent in choosing an appropriate reliability estimate. Here, we report ω_t coefficient, which by including the factor loadings in its formula, is more suitable and stable for reporting internal structure and reliability of multi–item scales since it corrects the underestimation bias of α when the assumption of tau equivalence is violated (Flora, 2020). In addition, different studies show that it is one of the best alternatives for estimating reliability (Revelle & Zinbarg, 2009; Zinbarg et al., 2006). Here, we report foundation-level ω_t coefficients across 19 populations (Table 5). As can be seen in Table 5, ω_t coefficients ranged between .73 and .95 (average ω_t coefficients: care = .90; equality = .89; proportionality = .83; loyalty = .89; authority = .86; and purity = .82). Hence, all six scores computed by averaging items for the six foundations are internally consistent across nations. Cronbach's α coefficients are available for comparison purposes on Supplemental Materials.

Measurement Invariance

The alignment method can be summarized in two steps (Asparouhov & Muthén, 2014; Fischer & Karl, 2019). First, an unconstrained configural model is fitted across all populations. To allow the estimation of all item loadings in the configural model, we fixed the factor means to 0 and the factor variances to 1. Second, we optimized the configural model using a component loss function with the aim of minimizing the noninvariance in factor means and factor variances for each group (for a detailed mathematical description, see Asparouhov & Muthén, 2014). This optimization process terminates at a point at which "there are few large noninvariant measurement parameters and many approximately noninvariant parameters rather than many medium-sized noninvariant measurement parameters" (Asparouhov & Muthén, 2014, p. 497). Overall, the alignment method allows for the estimation of reliable means despite the presence of some measurement noninvariance. Muthén and Asparouhov (2014) suggested a threshold of 25% noninvariance as acceptable. The resulting model exhibits the same level of fit as the original configural model but is substantially less noninvariant across all parameters considered. The percentage of noninvariant parameters in our invariance alignment method with post hoc item parameter constraints can be seen in Table 6. Effect sizes of approximate invariance based on R^2 have been proposed by Asparouhov and Muthén (2014). R^2 values of close to 1 suggest a greater degree of invariance, while values close to 0 indicate noninvariance (Asparouhov & Muthén, 2014). These are calculated separately for item loading and intercepts, presented in Table 6. As can be seen, all foundations except purity meet the threshold of 25% noninvariance, meaning that scores on care, equality, proportionality, loyalty, and authority can be reliably compared across cultural groups. For purity, caution should be practiced when comparing group-level means. In the present sample, the source of noninvariance in purity was mostly due to unique item intercepts in Argentina (six unique parameters; 5.3%) and Chile (four unique parameters; 3.5%). Among purity items, the item "I think the human body should be treated like a temple, housing something sacred within" was most noninvariant with 10 unique parameters (8.7%). Hence, this item may elicit different patterns of responding across different populations.

The Equality–Proportionality Link

One of the novel aspects of the present work is theoretically distinguishing between equality and proportionality. If the two constructs are distinct and psychometrically nonredundant, we should find only small-to-moderate correlations between them. We examined the correlation between equality and proportionality

Table 5
Omega Coefficients (ω_t) Across Foundations and Nations

Nation	Care	Equality	Proportionality	Loyalty	Authority	Purity
Argentina	0.90	0.92	0.77	0.83	0.83	0.82
Belgium	0.92	0.90	0.78	0.88	0.80	0.83
Chile	0.92	0.88	0.82	0.90	0.88	0.83
Colombia	0.87	0.89	0.81	0.90	0.86	0.82
Egypt	0.89	0.87	0.81	0.87	0.86	0.83
France	0.92	0.90	0.80	0.90	0.83	0.76
Ireland	0.92	0.86	0.86	0.90	0.90	0.85
Japan	0.88	0.89	0.83	0.89	0.84	0.73
Kenya	0.89	0.85	0.85	0.89	0.89	0.82
Mexico	0.91	0.87	0.83	0.86	0.86	0.80
Morocco	0.91	0.88	0.89	0.89	0.86	0.82
New Zealand	0.93	0.92	0.82	0.93	0.89	0.86
Nigeria	0.85	0.87	0.79	0.85	0.82	0.75
Peru	0.89	0.90	0.85	0.84	0.86	0.83
Russia	0.90	0.90	0.83	0.90	0.89	0.86
Saudi Arabia	0.89	0.89	0.87	0.88	0.84	0.76
South Africa	0.88	0.91	0.82	0.90	0.84	0.85
Switzerland	0.90	0.95	0.83	0.91	0.89	0.84
UAE	0.94	0.89	0.93	0.91	0.90	0.87
Range	0.85–0.94	0.85–0.95	0.77–0.93	0.83–0.93	0.80–0.90	0.73–0.87

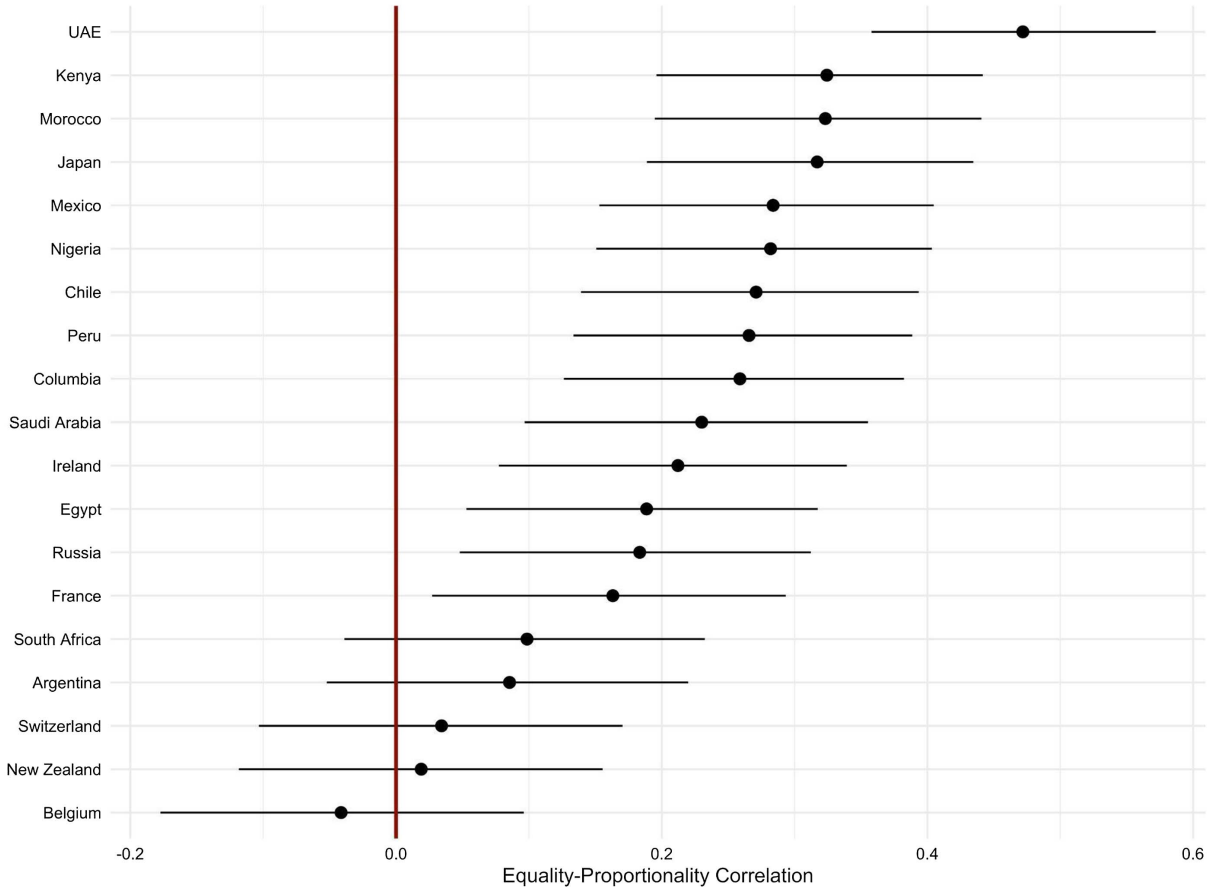
Table 6
The Measurement Invariance Alignment Results (Study 2)

Foundation	Loading R^2	Intercept R^2	% noninvariance item parameters (loadings)	% noninvariance item parameters (intercepts)
Care	.994	.999	0.0%	5.3%
Equality	.988	.995	0.0%	21.9%
Proportionality	.977	.999	0.0%	11.4%
Loyalty	.982	.998	0.0%	24.6%
Authority	.982	.996	0.0%	16.7%
Purity	.968	.989	2.6%	39.5%

Note. A threshold of 25% noninvariance or less is considered acceptable (Muthén & Asparouhov, 2014).

across all 19 populations and we found support for our prediction. Indeed, equality and proportionality were weakly positively correlated (average Pearson correlation coefficient = .21, *SD* = .13). Equality and proportionality were most related to one another in the UAE ($r = .47$, 95% CI [.36, .57], $p < .001$), while the smallest correlation was observed in Belgium ($r = -.04$, 95% CI [−.18, .10], $p = .556$). The correlations and their 95% CI are visually presented in Figure 2. Based on these findings, equality and proportionality may be considered orthogonal to one another, or only slightly positively correlated. WEIRDness was positively associated with orthogonality of equality and proportionality, $r = .40$, Conley *SE* = 0.32. For example, in nations such as New Zealand, Belgium, and Switzerland, people's scores on equality do not tell us anything about their concerns regarding merit and deservingness.

Figure 2
The Correlations Between Equality and Proportionality (Study 2)



Note. The error bars represent 95% confidence interval. The vertical line represents a zero correlation. See the online article for the color version of this figure.

Table 7
Means (and Standard Deviations) of Moral Foundations Across 19 Nations (Study 2)

Nation	Care	Equality	Proportionality	Loyalty	Authority	Purity
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Argentina	3.84 (0.77)	2.81 (1.01)	3.91 (0.66)	3.58 (0.82)	3.67 (0.73)	2.60 (0.82)
Belgium	3.91 (0.73)	3.20 (0.94)	3.91 (0.56)	3.62 (0.77)	3.70 (0.64)	3.01 (0.74)
Chile	3.77 (0.82)	2.77 (0.88)	3.70 (0.69)	3.45 (0.88)	3.67 (0.81)	2.54 (0.85)
Colombia	3.83 (0.71)	2.91 (0.90)	3.69 (0.68)	3.67 (0.82)	3.84 (0.68)	2.98 (0.86)
Egypt	4.38 (0.60)	3.56 (0.94)	4.37 (0.58)	4.42 (0.62)	4.18 (0.68)	4.19 (0.63)
France	4.08 (0.68)	3.23 (0.92)	4.12 (0.54)	3.86 (0.74)	3.88 (0.62)	3.09 (0.74)
Ireland	4.01 (0.79)	2.94 (0.93)	3.73 (0.77)	3.29 (0.98)	3.49 (0.91)	2.51 (0.93)
Japan	3.03 (0.77)	2.27 (0.78)	3.14 (0.73)	2.66 (0.82)	2.67 (0.66)	2.63 (0.69)
Kenya	4.2 (0.77)	2.88 (0.97)	3.78 (0.79)	3.95 (0.90)	4.07 (0.80)	3.58 (0.83)
Mexico	3.77 (0.79)	2.87 (0.91)	3.80 (0.70)	3.78 (0.75)	3.94 (0.67)	2.81 (0.81)
Morocco	4.21 (0.78)	3.36 (0.97)	4.18 (0.71)	4.16 (0.82)	3.95 (0.76)	3.93 (0.73)
New Zealand	3.84 (0.78)	2.61 (1.02)	3.61 (0.71)	3.22 (1.00)	3.48 (0.87)	2.58 (0.98)
Nigeria	4.32 (0.64)	2.90 (1.03)	4.14 (0.67)	4.11 (0.74)	4.21 (0.61)	3.80 (0.77)
Peru	3.62 (0.73)	2.63 (0.92)	3.75 (0.69)	3.73 (0.76)	3.81 (0.69)	3.00 (0.82)
Russia	3.96 (0.75)	3.24 (0.87)	4.27 (0.48)	3.87 (0.81)	3.68 (0.76)	3.25 (0.80)
Saudi Arabia	4.24 (0.75)	3.32 (0.93)	4.18 (0.69)	4.20 (0.78)	4.07 (0.73)	3.98 (0.72)
South Africa	4.21 (0.69)	3.01 (0.92)	4.03 (0.64)	3.85 (0.86)	4.00 (0.73)	3.40 (0.94)
Switzerland	3.95 (0.68)	3.27 (0.98)	3.84 (0.64)	3.58 (0.85)	3.52 (0.81)	2.95 (0.79)
UAE	4.01 (0.92)	3.28 (0.93)	3.96 (0.89)	4.02 (0.91)	3.91 (0.89)	3.74 (0.85)

Cross-Societal Differences

After measurement invariance was evidenced, we proceeded to examine cross-societal differences (see Table 7). Since, nations are nonindependent data points, we used Conley standard errors (Conley, 1999) to account for potential dependence based on spatial proximity in our data. Haversine distances were used to account for spatial autocorrelation. We then examined the relationship between WEIRDness cultural distance and moral foundations conditioned on participants' latitude and longitude. We also accounted for multiple comparisons by applying Bonferroni correction, which is used when several exploratory tests (here, six) are performed simultaneously. Cultural distance from the United States (less WEIRDness) was associated with higher scores on care ($r = .16$, Conley $SE = 0.04$, Bonferroni-corrected $p = .002$); equality ($r = .16$, Conley $SE = 0.06$, Bonferroni-corrected $p = .076$); proportionality ($r = .18$, Conley $SE = 0.04$, Bonferroni-corrected $p < .001$); loyalty ($r = .25$, Conley $SE = 0.04$, Bonferroni-corrected $p < .001$); authority ($r = .15$, Conley $SE = 0.04$, Bonferroni-corrected $p = .001$); and purity ($r = .43$, Conley $SE = 0.05$, Bonferroni-corrected $p < .001$). Purity was the most strongly related foundation to WEIRDness, with participants from less-WEIRD populations endorsing it substantially more strongly.

Gender Differences

In this section, we examined nationally variable gender differences in moral foundations. Notably, only 1.3% of our sample ($n = 50$) identified as nonbinary, hence we did not have adequate statistical power to explore this population, and only included participants identifying as either "woman" or "man." Based on the findings of Atari, Lai, and Dehghani (2020), we expected to find female-favoring scores on care and purity. We estimated a random-intercept model allowing populations to vary in gender differences in each of the foundations. For care, the fixed effect of gender was in line with our prediction, but was not statistically significant ($B = -0.03$, $SE = 0.024$, $p = .259$), indicating that prior findings regarding gender differences in care are smaller, possibly negligible, when measured using MFQ-2 rather than MFQ-1. This might also be attributable to some particular MFQ-1 items tapping into neighboring constructs such as compassion and nurturing tendencies, while MFQ-2 items are more focused on generic alleviation of pain and suffering. Women scored substantially higher than men on equality ($B = -0.16$, $SE = 0.03$, $p < .001$) and purity ($B = -0.09$, $SE = 0.026$, $p < .001$). Men, on the other hand, scored significantly higher than women on proportionality ($B = 0.09$, $SE = 0.022$, $p < .001$), loyalty ($B = 0.06$, $SE = 0.027$, $p = .038$), and authority ($B = 0.06$, $SE = 0.024$, $p = .009$).

Furthermore, we calculated Mahalanobis' D (95% CI based on 10,000 bootstrap iterations), which is a global (i.e., multivariate) measure of gender differences (Del Giudice, 2009, 2022). Notably, while univariate differences (e.g., Cohen's d) are important, they "may easily miss the forest for the trees" (Del Giudice, 2022, p. 8). Morality, according to MFT, is multidimensional, hence univariate gender differences calculated for individual foundations can be incomplete, uninformative, or even misleading (Atari, Lai, & Dehghani, 2020). Furthermore, the way in which gender differences in multiple moral foundations yield a global effect size depends on the sign and size of their intercorrelations. The most accurate metric for assessing global gender differences across several variables is Mahalanobis' D, which is the multivariate generalization of the well-known univariate effect size Cohen's d (Huberty, 2005). Mahalanobis' D can be interpreted as the distance between the centroids of men's and women's distributions across foundations, relative to the standard deviation along the axis that links these centroids (see Del Giudice, 2022; Eagly & Revelle, 2022).

Since D can overestimate gender differences in small samples and underestimate them when using unreliable measurements, we corrected for both of these biases by calculating a disattenuated, bias-corrected version of D, known as D_{cu} (Del Giudice, 2022). Multivariate gender differences in moral foundations (D_{cu}) were

smallest in France ($D = 0.461$, 95% CI [0.149, 0.610], $D_{cu} = 0.357$) and largest in Mexico ($D = 0.556$, 95% CI [0.225, 0.726], $D_{cu} = 2.130$). Across 19 populations, D_{cu} was large in size, $M = 1.06$, $Md = 0.92$, $SD = 0.55$. Larger D_{cu} values indicate more gender differentiation in overall pattern of moral judgments (Atari, Lai, & Dehghani, 2020). We examined the correlation coefficient between the WEIRDness cultural distance and D_{cu} and found that WEIRDer populations had slightly *smaller* multivariate gender differences in moral values ($r = .43$, Conley $SE = 0.17$, $p = .026$).

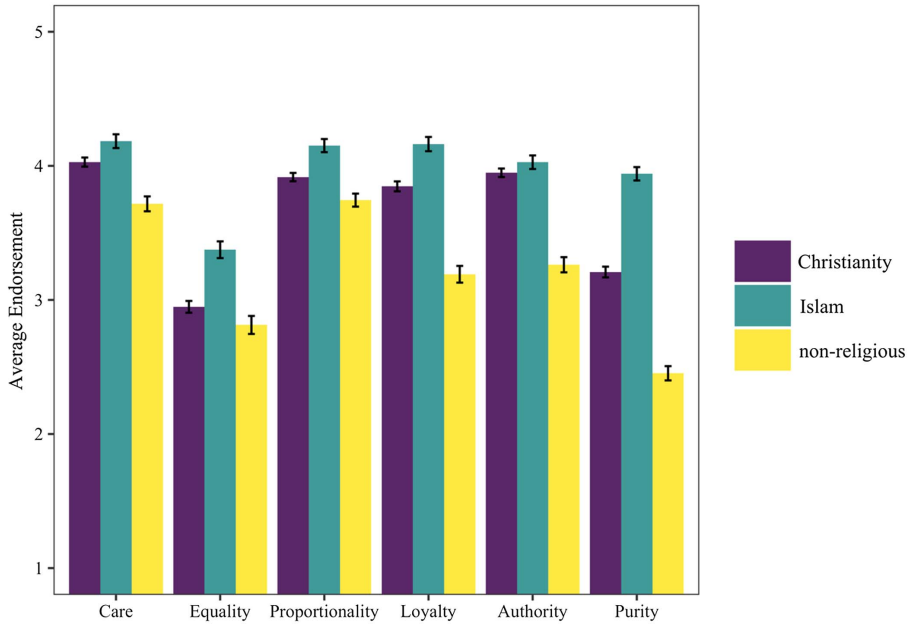
Religious Differences

We first examined moral foundations as a function of religious affiliation. Since we did not have enough data on individuals affiliating with Judaism ($n = 25$), Hinduism ($n = 14$), Buddhism ($n = 91$), and "other" affiliations ($n = 244$) in our data, we excluded these participants, leaving 3,527 individuals associating with Christianity ($n = 1803$), Islam ($n = 909$), and no religious affiliation ($n = 815$). One participant chose not to report their religious affiliation. For care, an analysis of variance suggested significant between-religion differences (Welch-corrected $F = 75.85$, $\omega^2 = 0.08$, $p < .001$), such that Muslims ($M = 4.18$) scored higher than nonreligious individuals ($M = 3.72$, Holm-corrected $p < .001$) and Christians ($M = 4.03$, Holm-corrected $p < .001$). For equality, there was a significant difference between religious affiliations (Welch-corrected $F = 85.03$, $\omega^2 = 0.09$, $p < .001$), such that Muslims ($M = 3.37$) scored higher than nonreligious individuals ($M = 2.81$, Holm-corrected $p < .001$) and Christians ($M = 2.95$, Holm-corrected $p < .001$). For proportionality, there was a significant difference between groups (Welch-corrected $F = 67.38$, $\omega^2 = 0.07$, $p < .001$) with Muslims ($M = 4.15$) scoring higher than nonreligious individuals ($M = 3.74$, Holm-corrected $p < .001$) and Christians ($M = 3.92$, Holm-corrected $p < .001$). For loyalty, there was a significant difference between groups (Welch-corrected $F = 274.62$, $\omega^2 = 0.24$, $p < .001$) with Muslims ($M = 4.16$) scoring higher than nonreligious individuals ($M = 3.19$, Holm-corrected $p < .001$) and Christians ($M = 3.85$, Holm-corrected $p < .001$). For authority, there was a significant difference between groups (Welch-corrected $F = 248.45$, $\omega^2 = 0.22$, $p < .001$) with Muslims ($M = 4.03$) scoring higher than nonreligious individuals ($M = 3.26$, Holm-corrected $p < .001$) and Christians ($M = 3.95$, Holm-corrected $p = .027$). Finally, for purity, there was a significant difference between groups (Welch-corrected $F = 799.96$, $\omega^2 = 0.46$, $p < .001$) with Muslims ($M = 3.94$) scoring higher than nonreligious individuals ($M = 2.45$, Holm-corrected $p < .001$) and Christians ($M = 3.21$, Holm-corrected $p < .001$). These differences are shown in Figure 3.

We then examined the relationship between religiosity and all six foundations using a cross-classified, random-intercept multilevel model wherein participants are nested in their nations (19 groups) as well as religions (three groups). Cross-classified multilevel modeling involves (at least) two multilevel data structures due to lower-level entities' double (or more) cluster memberships. Simulation studies have shown the undesirable consequences of mis-specifying a cross-classified structure when raw data are analyzed (Meyers & Beretvas, 2006). In addition, these models allow researchers to partition variance attributable to different cluster memberships. Here, our participants had multiple cluster memberships, which were themselves independent: country membership and religious affiliation. The results of the cross-classified, random-intercept multilevel model suggested that care ($B = -0.13$, $SE = 0.063$, $p = .043$) and proportionality ($B = -0.43$, $SE = 0.069$, $p < .001$) were negatively associated with

Figure 3

Endorsement of Moral Foundations Across Religious Affiliations (Study 2)



Note. Whiskers represent 95% confidence interval (CI). See the online article for the color version of this figure.

religiosity, while equality ($B = 0.12$, $SE = 0.043$, $p = .006$), loyalty ($B = 0.16$, $SE = 0.069$, $p = .019$), authority ($B = 0.20$, $SE = 0.080$, $p = .013$), and purity ($B = 0.75$, $SE = 0.060$, $p < .001$) were positively associated with religiosity. Cross-religion variation ($SD = 2.16$) was substantially larger than cross-nation variation ($SD = 0.36$).

Ideological Differences

We conducted a random-intercept, multilevel model to predict political ideology based on all six moral foundations. We found care ($B = -0.25$, $SE = 0.070$, $p < .001$) and equality ($B = -0.57$, $SE = 0.048$, $p < .001$) to be negatively correlated with political conservatism, while proportionality ($B = 0.30$, $SE = 0.075$, $p < .001$), loyalty ($B = 0.24$, $SE = 0.075$, $p = .001$), authority ($B = 0.55$, $SE = 0.087$, $p < .001$), and purity ($B = 0.13$, $SE = 0.063$, $p = .039$) were positively associated with right-wing ideology. That care is associated with liberal ideology, and that loyalty, authority, and purity are associated with conservative ideology are consistent with prior work (Graham et al., 2009; Kivikangas et al., 2021; McAdams et al., 2008). We also present novel findings with regard to the differential relationships between two novel foundations and political ideology. In line with our theorizing and prior work, we find that liberals are more concerned with equality and conservatives are more concerned with proportionality.

Next, on an exploratory basis, we examined the (non)linearity of the relationship between political ideology and moral foundations by testing the extent to which political conservatism's quadratic term (i.e., political ideology squared) predicts foundations above and beyond the linear effect described above. We conducted six random-intercept, multilevel models wherein political conservatism and its quadratic effect predicted each moral foundation separately. The quadratic effect of political ideology was nonsignificant for care ($B = -0.002$, $SE = 0.002$, $p = .275$), loyalty ($B = 0.0002$, $SE = 0.002$, $p = .916$), and authority ($B = 0.0003$, $SE = 0.002$, $p = .836$). It was also weak for purity ($B = 0.003$, $SE = 0.002$, $p = .039$). However, we found evidence for the polynomial relationship between equality, proportionality, and political ideology. For equality, the quadratic effect was positive (unlike its linear relationship; $B = -0.06$, $SE = 0.006$, $p < .001$), suggesting a U-shaped relationship where extremes score higher on equality ($B = 0.02$, $SE = 0.002$, $p < .001$). For proportionality, above and beyond its linear positive relationship ($B = 0.05$, $SE = 0.005$, $p < .001$), we found a negative quadratic effect ($B = -0.006$, $SE = 0.001$, $p < .001$), indicating an inverse U-shaped relationship between proportionality and political ideology. Polynomial relationships for all foundations across all countries (114 plots) are presented in Supplemental Materials.

Nomological Network of Foundations

We examined how the psychometric network of the six foundations looks across populations. Given some recent methodological reservations about higher order CFA (see Lee & Cadogan, 2013), we relied on EGA (Golino & Epskamp, 2017) to estimate the number of higher order dimensions in MFQ-2. Since equality and proportionality were not present in Graham et al. (2009), we performed community-detection analyses to investigate which moral foundations strongly cluster together. We used the "Walktrap" algorithm for community detection as it assigns nodes to a single cluster, has been demonstrated to yield reliable results (Pons & Latapy, 2006), and performs well on self-report data (Golino & Epskamp, 2017; Turner-Zwinkels et al., 2021). We ran the Walktrap algorithm via EGA. We estimated the Gaussian graphical model using graphical least absolute shrinkage and selection operator (Friedman et al., 2008) with extended Bayesian information criterion to select optimal regularization parameter. Similar to latent variable models (as in Graham et al., 2009), EGA effectively identifies the grouping of nodes (here, foundations) within a network; however, it either outperforms or is equal to other approaches used for estimating dimensions (e.g., parallel analysis, Kaiser–Guttman rule; Golino & Epskamp, 2017). Moreover, the network approach provides additional information about the relationships among the foundations while controlling for all possible relationships between pairs of foundations. Finally, since prior work shows that higher order networks of moral foundations may differ between populations (Atari, Graham, & Dehghani, 2020), we ran 19 different EGAs for the 19 populations we had data from.

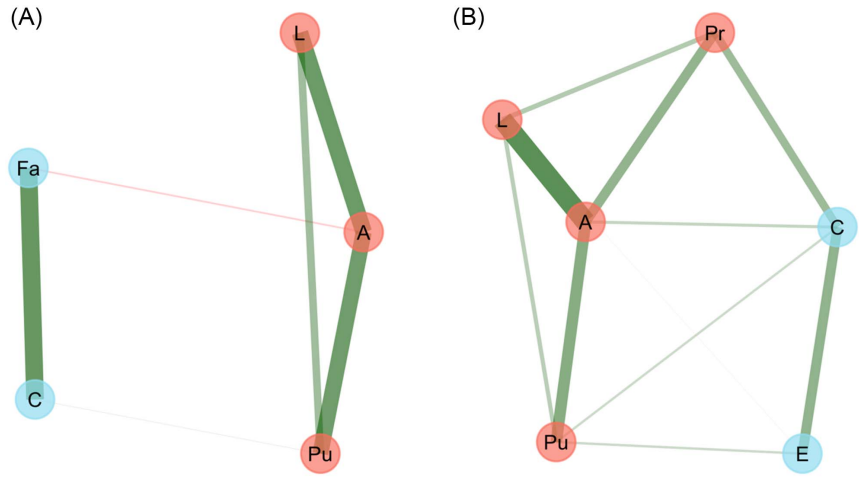
Before proceeding to the 19 networks, we conducted an EGA on U.S. data as a point of comparison using the MFQ-1 data on YourMorals ($N = 262{,}629$; Atari, Lai, & Dehghani, 2020) and using MFQ-2 data in Study 1c ($N = 515$). These two networks can be seen in Figure 4. As can be seen, using the five foundations (i.e., MFQ-1), there is a clear distinction between individualizing and binding values: care and fairness clearly form a subnetwork, almost completely segregated from loyalty, authority, and purity, which are intimately related themselves. The exploratory analysis of foundations using MFQ-2 (Study 1c) shows a similar pattern: care and equality are identified as a unique dimension, and loyalty, authority, proportionality, and purity form a second dimension. Next, we examine whether this clear distinction between individualizing and binding values is universal or WEIRD and U.S.-specific.

We then ran EGA for the 19 populations. All exploratory networks are presented in Figure 5. In all networks, γ and λ_{min} values were set to 0.5 and 0.1, respectively. The EGA analyses revealed one dimension in 16 samples and two dimensions in four samples (Ireland, New Zealand, and Peru). Our first take is that the individualizing-binding distinction may not be how moral foundations are organized universally; rather, the interrelations between the foundations should be considered culture dependent. These population-level differences are in line with the findings of Atari, Graham, and Dehghani (2020) and Turner-Zwinkels et al. (2021), demonstrating that moral foundations' network differs between groups. In the three nations in which we found a two-dimensional network, there was a somewhat consistent pattern. In Ireland and New Zealand, care and equality formed one dimension and the rest of the foundations formed a second dimension. In Peru, however, we found a dimension underlying equality and purity, while the rest of the foundations formed a second dimension. In all these models with two dimensions, the two subnetworks were moderately related to one another, and we found no evidence for complete segregation of these subnetworks. Accordingly, future research using MFQ-2 should be mindful of the cultural context when using higher order dimensions proposed by Graham et al. (2011) based on latent-variable models based on primarily North American and English-speaking participants.

We then investigated central nodes in the network of moral foundations across populations. We used a measure of node

Figure 4

The Network of Moral Foundations in the U.S. Samples Using MFQ-1 (Left) and MFQ-2 (Right)



Note. (A) The network of moral foundations in the United States using MFQ-1; (B) The network of moral foundations in the United States using MFQ-2. Different node colors denote belonging to different dimensions in exploratory graph analysis. Green edges denote a positive relationship and red edges indicate a negative relationship between two nodes after partial correlations with all other nodes have been taken into account. Width of edges indicate the strength of the unique relationship between two nodes. The network on the left (A) is based on YourMorals data ($N = 262{,}629$) and the one on the right (B) is based on Study 1c ($N = 515$). Fa = fairness; C = care; L = loyalty; A = authority; Pu = purity; E = equality; Pr = proportionality; MFQ = Moral Foundations Questionnaire. See the online article for the color version of this figure.

centrality that is considered most robust in the psychometric network literature, node strength, which denotes the sum of the weights connected to each node (Epskamp et al., 2018). Centrality indices related to each moral foundation across nations are shown in Figure 6. Notably, in 14 populations (73.7%), authority was the most central node among all foundations. Loyalty was the most central node in three samples (Saudi Arabia, Egypt, and Argentina). Finally, equality was the most central only in Belgium and proportionality was the most central only in Morocco. We then explored whether these variations could be related to WEIRD cultural distance: Care was more likely to be highly central in less-WEIRD societies (e.g., Nigeria, Kenya), $r(12) = .76$, 95% CI [.36, .92], $p = .002$. Loyalty was also more central in less-WEIRD societies, $r(12) = .62$, 95% CI [.13, .88], $p = .019$. Interestingly, despite the conceptual similarity between proportionality and equality, their relationship with WEIRDness diverged: in more WEIRD societies, equality was slightly more central, $r(12) = -.44$, 95% CI [−.80, .14], $p = .111$, whereas in less-WEIRD societies, proportionality was slightly more central, $r(12) = .45$, 95% CI [−.13, .80], $p = .104$.

Study 3

Study 3 was designed with three aims: (a) to establish the convergence of MFQ-2 scores with those of MFQ-1 (Graham et al., 2011), (b) to examine substantive relations with and capacity to predict criterion variables; and (c) to compare the predictive power of MFQ-2 and MFQ-1 in predicting the amount of variance in external scale scores. We selected three external scales as criterion variables for each foundation (see Measures section). As such, Study 3 provides evidence that MFQ-2 accurately quantifies its intended latent constructs (i.e., six moral foundations), shares theoretically pertinent associations with other constructs in moral foundations' nomological network, and does not capture confounding extraneous variables (Chester & Lasko, 2021).

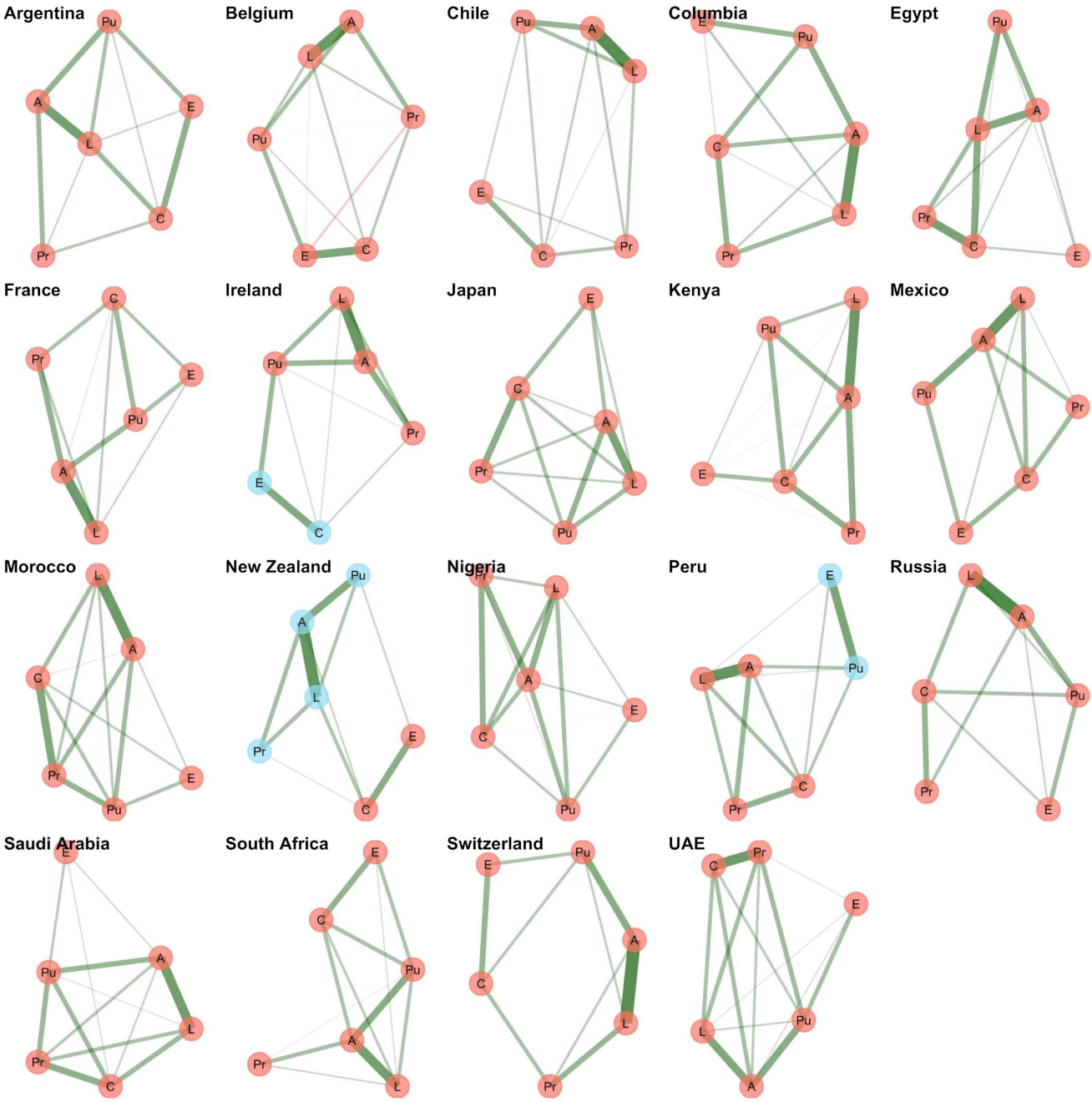
Method

Participants and Procedure

Since there were 18 criteria tested in this study, it was not practically feasible to have all participants complete all measures. Therefore, we collected six different samples, in which participants completed both MFQ-1 and MFQ-2 along with a battery of criterion scales, theorized to lie within moral foundations' nomological network. We aimed to collect a sample of 1,500 participants from the United States, India, and Canada on Cloud Research (Litman & Robinson, 2020). After removing participants who failed any of the three attention checks, 1,410 participants remained for analysis, mostly from the United States (82.1%). In terms of gender distribution, 642 participants identified as women, 762 identified as men, and six identified as nonbinary. Among American participants, most individuals identified as White (75.7%), followed by African American (13.3%) and Asian (5.4%). Based on our theoretical framework and prior research, we predicted 18 relationships (see Table 8). The measures we used across these six samples appear below. Our dependent variables,

Figure 5

Higher Order Networks Displaying the EGA-Identified Dimensions (Study 2)



Note. Different node colors denote belonging to different dimensions in exploratory graph analysis. Green edges denote a positive relationship and red edges indicate a negative relationship between two nodes after partial correlations with all other nodes have been taken into account. Width of edges indicate the strength of the unique relationship between two nodes. C = care; L = loyalty; A = authority; Pu = purity; E = equality; Pr = proportionality; EGA = exploratory graph analysis. See the online article for the color version of this figure.

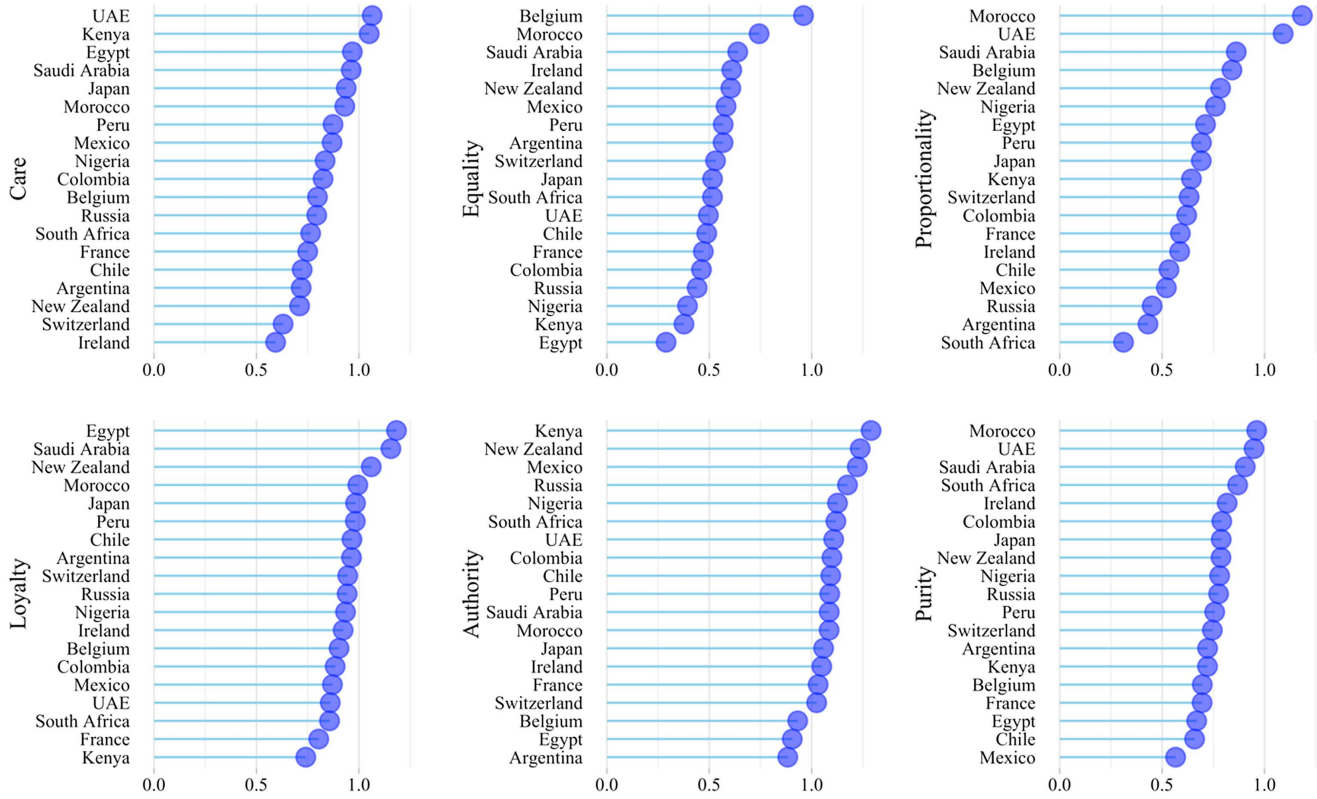
18 directional hypotheses, and sampling strategy were preregistered on OSF (https://osf.io/qfd93). This study was conducted in September 2021. This study was approved by the University of Southern California's IRB (UP-21-00635).

MFQ-2. We administered the 36-item MFQ-2 developed in Studies 1 and 2 (see Appendix). All 36 items were rated along a 5-point scale ranging from 1 (*does not describe me at all*) to 5 (*describes me extremely well*). In the present sample, α coefficients were .89, .87, .78, .85, .87, and .86 for care, equality, proportionality, loyalty, authority, and purity, respectively.

MFQ-1 (Graham et al., 2011). All participants completed the MFQ-1. Respondents rated the relevance items provided using

Figure 6

Centrality of Different Foundations Across Nations (Study 2)



Note. The *x*-axis represents the centrality of each foundation in the network of moral foundations. See the online article for the color version of this figure.

a 6-point scale from 0 (*not at all relevant*) to 5 (*extremely relevant*). The judgments items were rated along 0 (*strongly disagree*) to 5 (*strongly agree*). In the present sample, the internal consistency coefficients were .70, .67, .84, .81, and .87 for care, fairness, loyalty, authority, and purity, respectively.

Schwartz Values Survey (Schwartz, 1992). The Schwartz Values Survey (SVS) identifies 10 personal values. We report some of the SVS values that were previously used to examine the criterion validity of MFQ-1 by Graham et al. (2011; see Table 8). All items were rated from −1 (*opposed to my values*) to +5 (*of supreme importance*), where 0 indicates this value is "not important" for the person.

Interpersonal Reactivity Index (M. H. Davis, 1983). We used the Empathic Concern subscale of the Interpersonal Reactivity Index (IRI). Scores on this subscale are computed by averaging five items. This subscale measures other-oriented feelings of compassion for the misfortune of others (e.g., "I often have tender, concerned feelings for people less fortunate than me"). Items were rated along a 5-point scale ranging from 1 (*does not describe me well*) to 5 (*described me extremely well*). In the present sample, the internal consistency coefficient was .71.

Levenson Self-Report Psychopathy Scale (Levenson et al., 1995). The Levenson Self-Report Psychopathy Scale (LSRPS) was created to measure psychopathic traits and behaviors in the general population. This scale consists of 26 items rated along a 4-point Likert-type scale from 1 (*strongly disagree*) to 4 (*strongly agree*). The LSRPS was designed to reflect the dual-factor model of psychopathy, assessing primary psychopathy characterized by emotional deficits (e.g., lack of remorse) and manipulative behaviors (16 items), and secondary psychopathy, reflecting impulsivity, irresponsibility, and antisocial behaviors (10 items). An example item is "I enjoy manipulating other people's feelings." In the present sample, the overall internal consistency coefficient was .92.

Support for Redistribution Scale (Petersen et al., 2013). We used the six-item Support for Redistribution Scale (SRS) to measure participants' support for economic redistribution. All six items were rated along a 7-point scale from 1 (*strongly disagree*) to 7 (*strongly agree*). An example item was "The government should increase taxes and thus give more help to the poor." In the present sample, the internal consistency coefficient was .78.

Social Dominance Orientation (Ho et al., 2015). We used the extensively validated 16-item Social Dominance Orientation–7 (SDO-7) Scale (Ho et al., 2015), responding to items such as "An ideal society requires some groups to be on top and some to be on the bottom" (1 = *strongly oppose*, 7 = *strongly favor*). In the present sample, the internal consistency coefficient was .91.

Preference for the Merit Principle Scale (Davey et al., 1999). We used the Preference for the Merit Principle Scale (PMPS), which measures the extent to which people believe that

Table 8
The Correlation Coefficients Between MFQ-2 Scores and Criterion Variables

Criterion variables	Care	Equality	Proportionality	Loyalty	Authority	Purity
Empathic concern	**0.63*** **	−0.01	0.14*	−0.03	−0.09	−0.20**
SVS: Benevolence	**0.50*** **	0.13	0.35***	0.39***	0.48***	0.39***
Psychopathy	**−0.30*** **	0.32***	−0.06	0.21**	0.18**	0.43***
Support for redistribution	0.03	**0.56*** **	−0.10	0.05	−0.04	0.14*
Social dominance orientation	−0.36***	**−0.18** **	0.06	0.36***	0.40***	0.5***
SVS: Social justice and equality	0.51***	**0.29*** **	0.23***	0.11	0.09	0.12
SVS: Success	0.09	0.01	**0.22*** **	0.37***	0.42***	0.31***
Preference for the merit principle	0.26***	0.42***	**0.50*** **	0.50***	0.44***	0.47***
Belief in a just world	−0.03	0.14*	**0.29*** **	0.51***	0.53***	0.53***
SVS: Loyalty, national security, and family security	0.32***	−0.03	0.39***	**0.50*** **	0.58***	0.40***
Collectivism	0.40***	0.18**	0.39***	**0.59*** **	0.60***	0.42***
Group loyalty	0.02	0.03	0.34***	**0.78*** **	0.70***	0.60***
SVS: Tradition, obedience, social order, respect, and authority	0.12	0.15*	0.33***	0.71***	**0.76*** **	0.68***
Right-wing authoritarianism	−0.32***	−0.03	0.20**	0.61***	**0.69*** **	0.73***
Left-wing authoritarianism	−0.03	0.58***	0.02	0.10	**0.06**	0.30***
SVS: Clean, devout, spiritual, and self-discipline	0.21**	0.22***	0.25***	0.53***	0.60***	**0.72*** **
Disgust sensitivity	0.17*	0.23***	0.26***	0.30***	0.31***	**0.40*** **
Religiosity	0.01	0.18**	0.11	0.50***	0.53***	**0.72*** **

Note. SVS = Schwartz Values Survey; MFQ = Moral Foundations Questionnaire. Gray cells represent relationships for which we had a priori preregistered predictions. Bold correlations denote the correlation coefficients between MFQ-2 foundation scores and relevant criterion variables.
* $p < .05$. ** $p < .01$. *** $p < .001$.

outcomes and resources should be distributed based on merit (e.g., qualifications and achievements) rather than other factors such as need. An example item is "Qualifications ought to be given more weight than seniority when making promotion decisions." Items were rated along a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). In the present sample, the internal consistency coefficient was .70.

Belief in a Just World (Dalbert, 1999). We measured Belief in a Just World (BJW) with Dalbert's (1999) General (i.e., BJW-other) BJW subscales, which has six items. Items were rated along a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). An example item is "I am confident that justice always prevails over injustice." In the present sample, the internal consistency coefficient was .85.

Group Loyalty Scale (Beer & Watson, 2009). We measured group loyalty using the Group Loyalty Scale (GLS), which has eight items (e.g., "I would describe myself as a team player"). Items were rated along a 5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). In the present sample, the internal consistency coefficient was .92.

Individualism and Collectivism Scale (Triandis & Gelfand, 1998). Individualism and collectivism were measured using Triandis and Gelfand's (1998) scale. Participants rated the extent to which each of the 16 items described them. All items were rated along a 5-point scale ranging from 1 (*never*) to 5 (*always*). Four items measured vertical individualism (e.g., "It is important that I do my job better than others"), four measured horizontal individualism (e.g., "My personal identity, independent of others, is very important to me"), four measured vertical collectivism (e.g., "It is important to me that I respect the decisions made by my groups"), and four measured horizontal collectivism (e.g., "I feel good when I cooperate with others"). Here, we only report a composite collectivism score ($\alpha = .85$).

Right-Wing Authoritarianism (Altemeyer, 2006). The Right-Wing Authoritarianism (RWA) scale measures the degree to which people defer to established authorities, show aggression toward outgroups when sanctioned by authorities, and support traditional values endorsed by authorities. We used the most recent version of the RWA scale (Altemeyer, 2006) which has 22 items. Participants rated items (e.g., "Women should have to promise to obey their husbands when they get married") on a 9-point response scale ranging from 1 (*strongly disagree*) to 9 (*strongly agree*). In the present sample, the internal consistency coefficient was .94.

Disgust Scale–Revised (Olatunji et al., 2007). The Disgust Scale–Revised (DS-R) is a revised version of the 32-item Disgust Scale (Haidt et al., 1994). The DS-R consists of 25 items that measure how disgusting people find various stimuli. This scale consists of three subscales: contamination disgust, animal remainder disgust, and core disgust. In the first part of the measure, people indicate their agreement with items along a 5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). In the second part of the measure, participants indicate how disgusting an experience would be (1 = *not disgusting at all*; 5 = *extremely disgusting*). Here, we report an overall disgust sensitivity score ($\alpha = .86$).

Duke University Religion Index (Koenig et al., 1997). The Duke University Religion Index (DUREL) is a five-item measure developed for assessment of three main aspects of religiosity: organized religious activities (one item), nonorganizational religious activities (one item), and intrinsic religiosity (one item). The first two items are rated along a 6-point scale ranging from 1 (*never*) to 6 (*more than once a week/day*). The last three items, however, are rated along a 5-point Likert-type scale ranging from 1 (*definitely not true*) to 5 (*definitely true of me*). The DUREL's total scores can range between 5 and 27. In the present sample, the internal consistency coefficient was .92.

Left-Wing Authoritarianism (Costello et al., 2022). Left-Wing Authoritarianism (LWA) has been conceptualized as authoritarianism (e.g., aggression, submission, conventionalism) among individuals who oppose traditional established hierarchies of moral and practical authority. Despite right-wing authoritarianism receiving considerably more attention in the moral psychology literature, the conceptualization and measurement of LWA has only recently been done (Costello et al., 2022). We used the 39-item measure of LWA (e.g., "If I could remake society, I would put people who currently have the most privilege at the very bottom"). All items were rated along a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). In the present sample, the internal consistency coefficient was .95.

Demographics. At the end of the survey, all participants completed a set of demographic questions, including age, gender, education, religious affiliation, political ideology, and country of residence. All these questions were identical to those administered in Study 2.

Analytic Strategy

First, we examine the correlations between MFQ-2 scores and MFQ-1 scores using Pearson correlations. We also used a linear model to tease apart unique relationships between MFQ-1 foundation scores and MFQ-2 scores. That is, we ran six multivariate regression analyses: in each, all five MFQ-1 scores predicted one of the MFQ-2 scores as a dependent variable (i.e., care, equality, proportionality, loyalty, authority, and purity). Second, we examine the correlations between MFQ-2 foundation scores and the 18 criterion variables (three per foundation). Third, we broke down all external measure scores to their relevant subscale scores and used R^2 to quantify and compare the predictive power of both MFQ-2 and MFQ-1 in predicting these subscale scores. In this way, we examined how powerful MFQ-2 and MFQ-1 are in predicting related psychological variables.

Results

Convergence With MFQ-1

The correlation coefficients between MFQ-2 foundation scores and MFQ-1 foundations scores are summarized in Table 9. As can be seen, all foundations strongly relate to their predecessor subscale. In the case of fairness, it appears that MFQ-1s fairness captures both equality and proportionality, although its relationship to equality was stronger. This makes sense because some of the items in MFQ-1 directly tap into judgments about equality of outcomes (e.g., "I think it's morally wrong that rich children inherit a lot of money while poor children inherit nothing"). However, it is noteworthy that MFQ-1 fairness scores are moderately correlated with *both* equality and proportionality, *positively*. Of note, the correlation between equality and proportionality in the present sample was $r = .02$, 95% CI [−.03, .07], $p = .400$, consistent with the results of Study 2 wherein we found that in more WEIRD populations, these two constructs tend to be more orthogonal compared with less-WEIRD populations. Intercorrelations within the MFQ-1 and MFQ-2 are presented in Supplemental Materials.

Nomological Network

The correlation coefficients between MFQ-2 foundation scores and criterion variables are presented in Table 8 (for standardized regression coefficients, see Supplemental Materials). Out of our 18 predicted relationships, 17 were supported. The only correlation inconsistent with our predictions was between MFQ-2s authority and LWA ($r = .06$, 95% CI [−.07, .19], $p = .355$). The correlation did not change when we only examined U.S. participants ($r = .06$, 95% CI [−.08, .20], $p = .394$). Since this scale is mostly focused on antiauthority and antitradition sentiment (e.g., "Certain elements in our society must be made to pay for the violence of their ancestors"; see Costello et al., 2022), we predicted a negative relationship; however, we observed a positive, nonsignificant relationship. Other correlations supported the notion that MFQ-2s foundations have substantive relations with criterion variables.

People who score highly on MFQ-2's care show higher levels of empathic concern, take benevolence to be a guiding principle in their lives, and are less likely to have psychopathic traits. People who score highly on MFQ-2's equality show substantial support for redistributing resources in the society, have substantially less desire for some groups to be actively oppressed by others, have a stronger preference for intergroup equality, and consider social equality as a guiding principle in their life. People who score highly on MFQ-2's proportionality consider success as an important guiding principle to navigate their life, have a strong preference for merit, and believe that the world is generally a fair and orderly place wherein what happens to people is what they deserve. People who score highly on MFQ-2's loyalty tend to value nationality and loyalty, tend to meet the duties and obligations of one's social role to maintain group cohesion, and report to have remained loyal to their ingroup. People

Table 9
The Interrelationships Between MFQ-2 and MFQ-1 Scores

Score	MFQ-1-Care	MFQ-1-Fairness	MFQ-1-Loyalty	MFQ-1-Authority	MFQ-1-Purity
MFQ-2-Care	**.57***/.51***	.45***/.12***	−.02/−.08*	−.01/−.10*	.02/.04
MFQ-2-Equality	.25***/.04	**.33***/.29***	.19***/.28***	.09**/−.25***	.14***/.08*
MFQ-2-Proportionality	.20***/.08*	**.20***/.10****	.23***/−.07	.31***/.35***	.24***/−.01
MFQ-2-Loyalty	.13***/.04	.07**/−.07**	**.70***/.46***	.67***/.27***	.59***/.04
MFQ-2-Authority	.11***/.06*	.01/−.14***	.64***/.16***	**.70***/.46***	.63***/.17***
MFQ-2-Purity	.11***/−.01	.05/−.06*	.64***/.14***	.65***/.08*	**.76***/.60***

Note. Figures on the left side of the slash represent bivariate Pearson correlation and figures on the right side of the slash represent standardized regression coefficients in which all five MFQ-1 scores are accounted for. Bold correlations denote the correlation coefficient between MFQ-2 foundation scores and their relevant counterpart in MFQ-1. MFQ = Moral Foundations Questionnaire.
* $p < .05$. ** $p < .01$. *** $p < .001$.

who score highly on MFQ-2's authority tend to consider respect and obedience as important virtues as guiding principles and tend to value authoritarian submission, authoritarian aggression, and conventionalism. Finally, people who score highly on MFQ-2's purity tend to report higher levels of sensitivity toward disgusting things (e.g., animal remains, corpses, rotten food), value self-discipline and cleanliness, and report higher frequency of attending religious rituals, both organizationally (e.g., in a church), and nonorganizationally (e.g., saying prayers at home).

Predictive Power

We used both MFQ-1's and MFQ-2's scores in predicting subscale-level scores of all external measures. We collectively used 30 scores from SVS (self-transcendence, conservation, self-enhancement, and openness to change); LWA (antihierarchical aggression, anticonventionalism, and top-down censorship); empathic concern; group loyalty; LSRP (primary psychopathy and secondary psychopathy); BJW; DSR (core disgust, animal remainder, contamination); support for redistribution; Individualism and Collectivism Scale (ICS; horizontal individualism, vertical individualism, horizontal collectivism, and vertical collectivism); SDO (pro-dominance, con-dominance, pro-antiegalitarianism, con-antiegalitarianism); preference for the merit principle; RWA; DUREL (organizational religiosity, nonorganizational religiosity, and intrinsic religiosity); and political orientation. Across 30 regressions, MFQ-2 explained, on average, 37% of the variance in outcome variables (Md = 38%); however, MFQ-1 predicted, on average, 30% of the variance in all outcomes (Md = 26%; for a robustness check, see Supplemental Materials). The distribution of adjusted R^2 values and inferential statistics is presented in Figure 7. A paired t test indicated that MFQ-2 could explain significantly more variance in outcomes compared with MFQ-1 (t = 3.30, mean difference = 0.08, 95% CI [0.03, 0.12], p = .003, g_{Hedges} = .59).

Incremental Validity

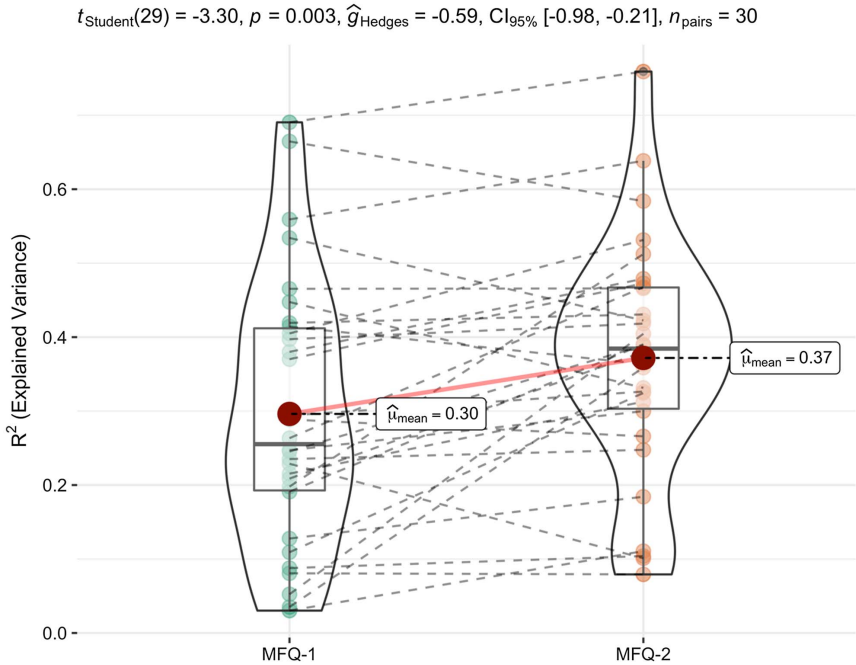
In order to examine the incremental validity of the MFQ-2, we tested improvement obtained by adding MFQ-2 scores to after accounting for all MFQ-1 scores in predicting 18 external measures detailed above. Across 18 measures, addition of MFQ-2 scores improved explained variance by 13.7%, on average, which was significantly higher than 0, 95% CI [8.17%, 19.30%], $t(17) = 5.21$, $p < .001$. These improvements are shown in Figure 8.

General Discussion

MFT (Graham et al., 2013; Haidt & Joseph, 2004) was developed by integrating evolutionary theories of human sociality and anthropological accounts of the breadth and variability of the moral domain (Fiske, 1992; Shweder et al., 1997). The original operationalization of MFT offered five moral foundations (care, fairness, loyalty, authority, and purity). For the past decade, the Moral Foundations Questionnaire (or MFQ-1) has been the primary tool with which these five foundations have been measured (Graham et al., 2011). Here, we revisited the assumptions and conceptualization of MFT and, based on data from 25 populations, we developed a new tool, the MFQ-2, which proves to be psychometrically superior across these cultural settings. This new instrument allowed us to demonstrate for the first time both the ubiquity of a set of specific

Figure 7

The Predictive Power of MFQ-1 and MFQ-2 in Predicting Outcomes



Note. MFQ = Moral Foundations Questionnaire; CI = confidence interval. See the online article for the color version of this figure.

moral concerns and the variability in the nomological network of these concerns across populations.

We had five major goals: (a) refining MFTs view on fairness by breaking it into equality and proportionality, and incorporating this theoretical refinement into the MFQ-2; (b) development and validation of MFQ-2 across cultures using local languages, and testing the structural validity and comparability of MFQ-2 scores across cultural contexts to make sure that MFQ-2 is truly a cross culturally meaningful and pragmatic tool; (c) examining how the network of moral foundations looks across populations and what foundations are more central depending on cultural context; (d) showing population-level and group differences (ideological, gender, and religious differences) using the novel MFQ-2; and (e) establishing external validity of the MFQ-2 by examining associations between criterion scales meant to capture relevant constructs.

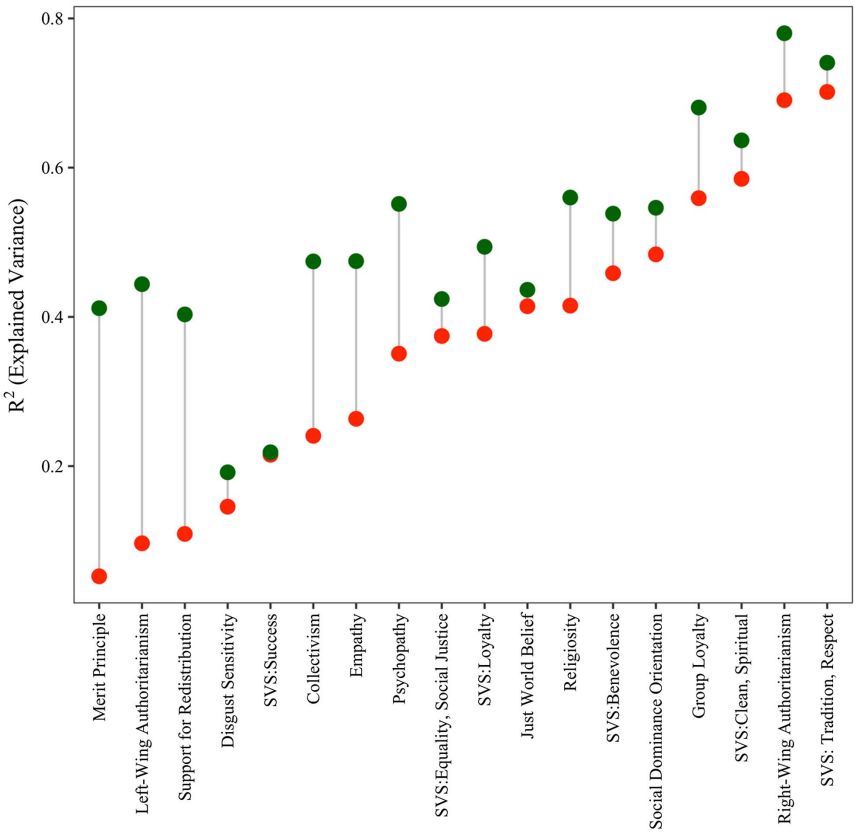
In three consecutive phases (cf. Flake et al., 2017), we report how the MFQ-2 fares in capturing the moral domain. We aimed to have six items per subscale, as is the case in MFQ-1 (Graham et al., 2011). In Studies 1a, 1b, and 1c, we compiled a 50-item pool based on data from diverse cultural backgrounds (India, Iran, Ecuador, China, United States). Notably, we believe it is crucial for a true non-WEIRD science of morality to start from non-WEIRD contexts in order to make sure that our measurements are not tuned to culturally unusual characteristics of certain populations. This approach, though encouraged by the theoretical roots of MFT (Shweder & Haidt, 1993), has remained ignored mostly because of lack of easy access (or even expensive access) to non-WEIRD populations in psychology (see Moshontz et al., 2018). In Study 2, we diversified our samples even more by recruiting nationally stratified by key demographic characteristics from 19 populations, most of which remain understudied in social and personality psychology (Thalmayer et al., 2021). We test structural validity of MFQ-2, its measurement invariance, and group differences in endorsement of moral foundations across these 19 populations. In Study 3, we examine how moral foundations, measured using MFQ-2, relate to relevant constructs, and show that the MFQ-2 outperforms the MFQ-1, and has substantial incremental validity, in predicting these criterion variables.

Moral Pluralism: Moral Systems as Networks

Throughout the history of moral psychology, various theorists have taken a monist approach, arguing that all of morality is based

Figure 8

The Incremental Validity of MFQ-2 in Predicting External Measures



Note. Red dots represent explained variance in a model with MFQ-1 scores as predictors. Green dots represent explained variance in a model with all MFQ-2 and MFQ-1 scores as predictors. SVS = Schwartz Values Survey; MFQ = Moral Foundations Questionnaire. See the online article for the color version of this figure.

on or can be reduced to a single construct or virtue. For Kohlberg (1969), for example, it was justice. For Baumard et al. (2013), it is mutualistic fairness. For Gray et al. (2014; Schein & Gray, 2015; Schein & Gray, 2018), all morality is harm. All these views were put forward in WEIRD populations and by WEIRD researchers (although Kohlberg also conducted cross-cultural research in developmental psychology; see Snarey et al., 1985). Other than "Occam's razor" (a principle of theory construction suggesting that, other things equal, explanations that posit fewer entities are to be preferred to explanations that posit more), these theories do not make explicit why humans should have one and only one morality. We have previously argued for the advantages of moral pluralism over moral monism, in terms of both the preponderance of scientific evidence (Graham, 2015; Graham et al., 2018) and the usefulness of moral pluralism to the interdisciplinary study of morality in general (Graham et al., 2013; Haidt, 2007). Different cognitive phenotypes contained in different moral foundations are likely cultural "kludges" (Stich, 2006) or cultural adaptations to particular socioecological conditions. The diverse plurality of morality makes complex forms of cooperation and sociality possible in the diverse culturally constructed worlds we live in Greene (2013). The cross-societal variations shown here—not only in mean levels of moral foundation endorsements, but in the very nomological networks in which those moral concerns relate to one another—further demonstrate the discoveries made possible by a pluralistic approach to moral judgments and concerns.

A methodological challenge for these alternative theories of morality has been to examine the position of the proposed values within the structure of the moral system as a whole. Here, as a solution to this challenge, we used networks of moral foundations wherein interrelationships between foundations are directly modeled as a network of interacting nodes, allowing to locate "central" nodes in the network in addition to other advantages. A network approach is particularly appropriate for a pluralistic view on morality as it accommodates many components within the same analysis in which multiple foundations, as well as their interrelationships, can be examined simultaneously (e.g., Brandt et al., 2019). Using methodological advances from psychometric network analysis, we quantified central and peripheral foundations across populations, as well as the relationship between foundation-level centrality and the cultural WEIRDness continuum.

Less-WEIRD Morality

While Graham et al. (2011) contended that "one does not need to travel to non-Western nations to find [MFT's] broader conception of morality" (p. 380), one certainly needs to collect high-quality data from non-Western nations, and collaborate with non-Western researchers (Medin et al., 2010), to ascertain that moral psychological theories hold firmly across various human populations, not just a small slice of human diversity. This was our motivation in recruiting a diverse group of participants across our studies. MFT was created as an evolutionarily informed cultural theory of human morality; hence, it is imperative that its claims be tested in WEIRDer and less-WEIRD populations and languages.

Breaking down fairness into equality and proportionality is one step toward better understanding fairness and justice concerns among populations. For example, Starmans et al. (2017) posited that "outside of the United States and Europe … there are wide differences in fairness concerns across world cultures" (p. 3), concluding that the distinct preferences for both equality and proportional outcomes are predominant in many cultures. Interestingly, we found that in less-WEIRD populations, proportionality is more likely to be an influential node in the nomological network of moral foundations, whereas in WEIRD populations, equality was found to be more central.

Notably, only recently has it become possible to test the relationship between psychological constructs and a continuous measure of WEIRDness empirically, especially with the advent of the WEIRDness cultural distance (Muthukrishna et al., 2020). While many researchers have speculated about non-WEIRD moral concerns, and some researchers having erroneously dichotomized the WEIRD spectrum (e.g., Doğruyol et al., 2019; see Apicella et al., 2020, detailing why this is a bad idea), no study to our knowledge had examined the relationship between WEIRDness and moral foundations. In the present research, we found that culture-level endorsements of purity and loyalty are higher in less-WEIRD populations. Therefore, purity and loyalty may be considered least WEIRD of the moral foundations, being substantially more salient in nations such as Egypt, Saudi Arabia, and Morocco. Although future research is encouraged to replicate these findings in larger samples from more cultures, including traditional small-scale communities (e.g., Purzycki et al., 2018).

One of the implications of the present research is its application in understanding and assessing less-WEIRD morality. We achieve this by two means: first by widening our top-down theoretical lens, which better captures less-WEIRD conceptions of morality (see Willard et al., 2020), particularly fairness; and second, by diversifying our samples using which we developed MFQ-2 (see Apicella et al., 2020; Henrich, 2020). In addition, using Muthukrishna et al.'s (2020) newly validated index of WEIRDness cultural distance, we tested novel predictions about different moral foundations in less-WEIRD cultures, finding that purity and loyalty are particularly higher in less-WEIRD populations such as Egypt and Saudi Arabia. Our approach has important implications for moral psychological research because moral cognition may be more like a kludge, shaped by local social norms and socioecological factors (e.g., Atari et al., 2022) and other cognitive processes (e.g., Khoudary et al., 2022; Pennycook et al., 2014) than a single cognitive architecture (Stich, 2006); hence, it is imperative that our tools are created with this human diversity in mind, making sure that our tools are understandable and usable across less-WEIRD populations. In addition to collecting data from many different populations, we also maximized, as much as possible, religious diversity in our sample. Most of existing research on the relationship between religious beliefs and morality has focused on Christianity (Bloom, 2012; Norenzayan, 2016; White et al., 2019). This focus on Christianity, and even more narrowly, Protestantism, has been argued to be a common feature of existing research on psychology of religion, as previously highlighted by cross-cultural scholars of religion (e.g., Apicella et al., 2020; Saroglou & Cohen, 2013). Tapping into the religious diversity across nations, we tested the relationships between moral foundations and religious identity, as well as the strength of religious practice.

Our results also have some implications for the conceptual clarity of the purity foundation. MFTs purity foundation has been criticized for including notions of religion as an inherent feature of purity, which could lead to a number of conceptual issues (Crone, 2022). After all, if purity is simply what God disapproves of, then purity can arguably be regarded as a metaethical concern such that whatever

God disapproves of is "impure," which could include a diverse array of transgressions such as eating pork, charging interest, protesting against clergy, or neglecting to fight for one's religion. This conceptual problem was further exacerbated since MFQ-1's purity subscale included an item that included the word "God" in it (i.e., "Whether or not someone acted in a way that God would approve of"), which served as a proxy for religious identity (Crone & Laham, 2022). In developing the MFQ-2, we carefully created items that do not serve as a proxy for religiosity. Still, we believe and demonstrate that purity and religiosity are highly correlated, but correlations based on the MFQ-2 are free of the confounds that MFQ-1 introduced. While we sought to maximize conceptual coverage by including items capturing purity concerns in different domains (e.g., foul language, sexuality, natural medicines), construct validity analyses showed that purity remained a unified construct. This advance provides the opportunity to more reliably measure this less-WEIRD moral foundation and clarifies some of the murkiness in the conceptualization of purity (see Crone, 2022; Graham et al., in press; Gray et al., 2022).

Differences (and Also Similarities) Across Cultures

Using the MFQ-2 in Study 2, we discovered three group differences in moral foundations: gender differences (see Atari, Graham, & Dehghani, 2020), religious differences (see Graham & Haidt, 2010), and ideological differences (see Kivikangas et al., 2021). Our examination of cross nationally variable gender differences suggested that women cared more about equality and purity than men. Men on the other hand scored slightly higher than women on loyalty, authority, and proportionality. Women's stronger emphasis on equality and purity may be related to their parental care systems and disgust sensitivity, previously researched in evolutionary social sciences (Al-Shawaf et al., 2018; Benenson et al., 2022; Rozin et al., 2000). These gender differences are consistent with prior work showing that women attribute more importance to understanding, appreciation, tolerance, and protection of the welfare of all humans and for nature across populations (Schwartz & Rubel, 2005). Relatively small gender differences in loyalty and authority (i.e., small in size and variable across populations) are consistent with Atari, Lai, and Dehghani (2020) and suggest that motivations for ingroup loyalty and support for hierarchical social structures are not substantially different between women and men. This finding is in line with evolutionary anthropological work investigating gender differences in political leadership in small-scale societies, demonstrating that gender differences in leadership activities and coordination of ingroup members are not a direct product of differences in motivation for status and leadership, rather an indirect product of gender differences in schooling, cooperation strategies, and sexual division of labor (Von Rueden et al., 2018).

With regard to religious differences, we found that more religious individuals tend to score lower on care and proportionality, while being more likely to score higher on loyalty, authority, purity, and equality. These strong associations between religious affiliation, religious practices, and endorsement of moral foundations are consistent with Graham and Haidt's (2010) argument that beliefs, rituals, and other facets of religious practice are best understood as means of creating a moral community. We propose, based on the present cross-societal findings, that this preference is best understood as emotive for an "egalitarian moral community" rather than a merit-based cooperative community. We found that religious differences exceed national differences in moral values, indicating that individuals who share a particular religious affiliation and level of commitment to religious practices are morally similar, both within and across countries (White et al., 2021).

With regard to ideological differences, we replicated the principal findings of Graham et al. (2009) and the meta-analytic results of Kivikangas et al. (2021). In particular, we found that conservatives tend to score higher on loyalty, authority, purity, and proportionality while scoring lower on care and equality. Our results are generally consistent with Kivikangas et al. (2021) who found that, with a few exceptions in their meta-analysis, care and fairness negatively, and loyalty, authority, and purity, positively correlate with right-wing political ideology. Indeed, prior MFT research did not have the equality–proportionality distinction. We find that liberals tend to value equality while conservatives tend to prioritize proportionality. These new findings are consistent with prior work finding that individuals on the right are more likely to endorse rewarding and punishing people on their merits (Arts & Gelissen, 2001; Haidt, 2012), while liberals are more likely to be egalitarian on different personality measures (Jost et al., 2003). The MFQ-2 provides the opportunity for future research to examine the diverging roles of equality and proportionality on an array of ideology-related outcomes.

Equality and Proportionality as Distinct Paths to Understanding Fairness

One of our theoretical revisions in this work is revisiting the concept of fairness in light of recent empirical findings. We break down fairness to more narrowly defined constructs in order to sharpen MFT's view on fairness. We defined equality in terms of a motive for balanced reciprocity, equal treatment, equal say, and equal outcome. Proportionality, on the other hand, is a psychological mechanism concerned with rewards and punishments to be proportionate to merit and deservingness, and benefits to be calibrated to the amount of contribution.

In our scale development procedure, we made sure that (a) items representing these two constructs were not Eurocentric (achieved by recursively soliciting feedback from a diverse group of social and personality psychologists; see Medin et al., 2010); and (b) items were not written with a particular political tone, which may inflate foundations' correlation with political ideology (e.g., some MFQ-1 fairness items have been shown to be particularly relevant in the American political context, which may have contributed to especially strong correlations between foundation scores and political ideology; Kivikangas et al., 2021).

The addition of foundations should come as no surprise; MFT theorists have explicitly welcomed new foundations to be appended to their framework as methods and theory codevelop in moral psychology. Particularly, with regard to addition of new foundations, Graham et al. (2013, p. 58) paraphrased Isaiah Berlin in writing that they "do not know how many moral foundations there really are. There may be 74, or perhaps 122, or 27, or maybe only five, but certainly more than one." Graham et al. (2011) posited that what their map of the moral domain originally offered (the five foundations) was "surely incomplete" (p. 382). These authors proposed that their empirical support for the theory was a good initial map of the major moral continents; however, "it is quite

possible that later research, using different items or different methods, would reveal that one of these continents is, like Eurasia, really two continents" (Graham et al., 2011, p. 382). That is exactly what we have found and proposed in the current work, taking one more step toward clarifying the nature and structure of the moral domain using a cultural psychological lens. This can open doors to many future investigations and novel theoretical questions. This proposition is a direct response to Graham et al. (2011) speculation that "whether a single foundation underlies intuitions about equality of opportunities and those about equality of outcomes [remains an open question]" (p. 382).

Indeed, both equality and proportionality have strong empirical evidence to warrant their consideration as separate moral foundations. Specifically, both are common in third-party normative judgments. Equality motivates people to be more sensitive to their relative payoffs, compared with others, rather than to the total amount they get (Bazerman et al., 1995). Moreover, individuals use either equality or proportionality heuristics to determine fair allocations among groups (Camerer & Thaler, 1995) and are sensitive to contextual information judgments about equality and proportionality (Andrejević et al., 2020). People have quick, affective reactions to both equality and proportionality transgressions (Sunar et al., 2021). Equality and proportionality concerns are culturally widespread as evidenced in the current work as well as cross-cultural and ethnographic work (Almås et al., 2010; Fiske, 1990; Whitehead, 2000). Both show up at young ages (Zhang, 2020), have been observed to some extent in nonhuman primates (Brosnan, 2013; see Chudek & Henrich, 2011), and are evolutionarily stable strategies for cooperation (Rai & Fiske, 2011). In sum, both equality and proportionality have good empirical evidence supporting our suggestion that they be considered separate moral foundations upon which societies build different fairness-related norms (e.g., eye-for-an-eye revenge norms), narratives (e.g., hero's journey), and institutions (e.g., court systems).

Our reconsideration of fairness judgments, implemented in the MFQ-2, can aid in our understanding of the current American culture war over fairness Hunter (1991), in which the left is concerned about justifying social inequalities and systemic racial inequality in the name of merit (e.g., Goudarzi et al., 2020), while the right often objects to disregarding one's talent and effort in the name of equality. Future studies investigating justice beliefs could benefit from considering how individual differences in equality and proportionality predict how people react to specific and culture-specific instances of injustice. For example, natural disasters and illness may be perceived to threaten principles of social equality, leading to compensatory action (Hafer & Rubel, 2015); however, these same experiences may seem morally justifiable when cultural narratives attribute them to notions of deservingness (Goudarzi et al., 2020; Sandel, 2020; Yan et al., 2023).

Nomological Network of Moral Foundations

Our findings provided compelling evidence that MFQ-2 captures more variance in a variety of outcomes compared with MFQ-1 across three populations. This is noteworthy given that MFQ-1 is already regarded as a powerful tool in predicting a wide array of outcomes ranging from political behavior (Kivikangas et al., 2021) to real-world hate group activities (Hoover et al., 2021). Even when we completely excluded proportionality, MFQ-2 still significantly outperformed MFQ-1, indicating that MFQ-2's superior predictive performance is not due to having several more items or a new subscale. This finding is promising as it opens the door to future theory-driven examination of morally relevant behaviors and judgments, as well as modeling approaches that use MFT to minimize out-of-sample prediction error in predicting a behavioral outcome (e.g., Reimer et al., 2022). Furthermore, Studies 2 and 3 collectively provided evidence that the individualizing-binding distinction made in Graham et al. (2011) may actually be culture dependent. Accordingly, one may not assume that two-dimensional higher order structure exists in all cultural contexts. This is a new insight into MFT, which is plausible since most of Graham et al.'s (2011) data were based on North American and English-speaking participants. The segregation of moral values into entirely isolated islands (individualizing and binding) appears to be a WEIRD phenomenon, rather than a universal feature of the moral domain.

Our network approach adds to another emerging line of work indicating that moral foundations are interconnected in different ways in different cultures (Atari, Graham, & Dehghani, 2020; Turner-Zwinkels et al., 2021). What causes the moral domain to separate into isolated islands that move away from one another? Future research is encouraged to investigate cultural-evolutionary processes that give rise to this segregation across time and space.

Study 3 further demonstrated that the six foundations were related to theoretically relevant constructs in predictable ways. The only exception was LWA, which yielded a nonsignificant correlation with authority. Interestingly, LWA was also unrelated to loyalty (on which conservatives tend to score higher) and care (on which liberals tend to score slightly higher); however, LWA was strongly associated with equality. These results suggest that, at least in the current framework of MFT, LWA may be seen as a form of intense egalitarianism, of the kind that has been visible in revolutions from the French revolution through the Bolshevik and Chinese cultural revolutions (Stone, 1980). It may have no relationship with the concerns for order and stability that are at the heart of the authority foundation.

MFQ-2

In the past few years, MFQ-1 has been rightly subjected to psychometric criticism regarding its structural validity, as well as internal consistency, especially in diverse, non-Western samples (e.g., D. E. Davis et al., 2017; Harper & Rhodes, 2021; Iurino & Saucier, 2020). In many of these studies, the original factor structure was not replicated, and foundation-level internal consistency coefficients were lower than conventional thresholds. These criticisms pointed to the need for a psychometrically superior and truly cross-cultural and cross-linguistic instrument, particularly because poor measurement qualities of common measures in social and personality psychology are central culprits in the replication crisis (which has sometimes been referred to as the measurement crisis; Flake & Fried, 2020).

In the entirety of the process of item reduction, we avoided relying on a single population to avoid cultural biases shaping the final battery of items in any form. The final 36-item MFQ-2 was developed with a diverse set of participants (Henrich et al., 2010) and by a diverse set of researchers and collaborators (Medin et al., 2010). We also employed

different methodological strategies, each of which has its own benefits and limitations. This multimethodological approach pushes against biases and inclinations inherent in particular methodological choices. For example, ESEM balances the advantages and disadvantages of EFA and CFA, item response theory-based methods such as the alignment method alleviate concerns about CFA-based methods in testing measurement invariance across many groups, and network psychometrics is a helpful toolbox to complement classical test theory (Golino et al., 2020). In sum, the MFQ-2 has desirable psychometric properties across almost all of the nations from which we had data in the current research. MFQ-2 scores also proved to be meaningfully comparable across cultures as measurement invariance was evidenced.

Limitations and Future Directions

The present research had some limitations that suggest important future directions. One such limitation is that we currently do not have cross culturally valid measures of other "candidate foundations," which have been proposed as potential moral foundations using the foundationhood criteria set by Graham et al. (2013) but have not gained consensus among researchers as foundation. Notable candidates are liberty (Iyer et al., 2012), honor (Atari, Graham, & Dehghani, 2020), honesty, ownership, and efficiency (see Graham et al., 2013). Other scholars have also built upon MFT, developing different typologies of moral values with slightly different lists of foundations. Curry et al. (2019), for example, have proposed seven moral foundations, including family, group, reciprocity, bravery, respect, fairness, and property, as part of their interesting work in this literature. Yet, Curry et al. (2019) left out purity as an important and less-WEIRD moral foundation. Our six-dimensional model is the most parsimonious model that captures the moral domain based on the current state of the art, and MFQ-2 is shown to be the best existing tool with which these moral intuitions can be measured. However, the addition of foundations—and the development of additional scales to measure those foundations—is a great next step for a pluralistic approach to human morality.

Second, while we collected data from 25 populations and seven languages, the present results are still based on a subset of these populations who were educated enough to complete the surveys online. Our sample did not include people from traditional, small-scale communities, whose means of living are subsistence based with everyday social interactions being mainly with local familiars such as their kinship network (e.g., Purzycki et al., 2018). To our knowledge, the present work is among the first to revise a commonly used psychometric measure in less-WEIRD populations, and the future work is encouraged to further examine our model in ethnographic work, cross-cultural research, and intersectional studies.

Moreover, one of the assumptions of general linear models is the independence of residuals—an assumption which is typically violated when using geographic data such as ours. Closer countries (e.g., Belgium and France) tend to be more similar to one another and more distant countries tend to be more dissimilar (e.g., Belgium and Saudi Arabia), resulting in higher false-positive rates (Ebert et al., 2022; Ward & Gleditsch, 2008). Our sample of countries was too small to conduct a formal test of nonindependence, but future research is recommended to perform spatial regressions that account for geographical nonindependence of countries.

Future research can extend other MFT-based measurement tools. Among others, the Moral Foundations Dictionary (Graham et al., 2009), Moral Foundations Dictionary 2.0 (Frimer et al., 2019), Moral Foundations Vignettes (Clifford et al., 2015), Moral Foundations Sacredness Scale (Graham & Haidt, 2012), Moral Foundations Twitter Corpus (Hoover et al., 2020), MapYourMorals (Hoover et al., 2021), the Socio-Moral Image Database (Crone et al., 2018), and Moral and Affective Film Set (McCurrie et al., 2018) can be updated in accordance with the new findings and refinements reported here, further generating testable hypotheses about human morality in different contexts which can be measured using different methodologies.

Conclusion

MFT was created in the early 2000s, a decade in which it still seemed possible that all populations would eventually become liberal democracies and all people would become somewhat WEIRD (e.g., Fukuyama, 2006). Now, in the 2020s, the future looks more morally diverse, politically chaotic, and eternally conflictual. Since the MFQ-1 was first published in 2011 the world has seen an increase in illiberal democracies and authoritarian states, the further fracturing of the "World Wide Web" into several state-run webs (e.g., in China and Iran), the migration of political discourse onto advertising-driven, algorithmically curated outrage platforms, and existentially threatening levels of partisan conflict in the United States. If ever there was a time when social scientists needed good tools for studying the values, judgments, and passions of diverse moral communities, it is now. We offer the MFQ-2 as a tool for our time.

References

Adams, J. S. (1963). Toward an understanding of equity. *Journal of Abnormal Psychology*, *67*(5), 422–436. https://doi.org/10.1037/h0040968

Adams, J. S. (1965). Inequity in social exchange. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 267–299). Academic Press.

Akhtar, N., Francis, L. J., Village, A., Sailer, A. B., Hasan, S. S., & McKenna, U. (2023). Testing the Moral Foundations Questionnaire within a Muslim society: A study among young adults in Pakistan. *Journal of Religious Education*, *7*(3), 1–18. https://doi.org/10.1007/s40839-023-00195-z

Al-Shawaf, L., Lewis, D. M., & Buss, D. M. (2018). Sex differences in disgust: Why are women more easily disgusted than men? *Emotion Review*, *10*(2), 149–160. https://doi.org/10.1177/1754073917709940

Almås, I., Cappelen, A. W., Sørensen, E. Ø., & Tungodden, B. (2010). Fairness and the development of inequality acceptance. *Science*, *328*(5982), 1176–1178. https://doi.org/10.1126/science.1187300

Altemeyer, B. (2006). *The authoritarians* [Unpublished manuscript]. Department of Psychology, University of Manitoba.

Amin, A. B., Bednarczyk, R. A., Ray, C. E., Melchiori, K. J., Graham, J., Huntsinger, J. R., & Omer, S. B. (2017). Association of moral values with vaccine hesitancy. *Nature Human Behaviour*, *1*(12), 873–880. https://doi.org/10.1038/s41562-017-0256-5

Andrejević, M., Feuerriegel, D., Turner, W., Laham, S., & Bode, S. (2020). Moral judgements of fairness-related actions are flexibly updated to account for contextual information. *Scientific Reports*, *10*(1), Article 17828. https://doi.org/10.1038/s41598-020-74975-0

Apicella, C., Norenzayan, A., & Henrich, J. (2020). Beyond WEIRD: A review of the last decade and a look ahead to the global laboratory of the future. *Evolution and Human Behavior*, *41*(5), 319–329. https://doi.org/10.1016/j.evolhumbehav.2020.07.015

Arts, W., & Gelissen, J. (2001). Welfare states, solidarity and justice principles: Does the type really matter? *Acta Sociologica*, *44*(4), 283–299. https://doi.org/10.1177/000169930104400401

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*(3), 397–438. https://doi.org/10.1080/10705510903008204

Asparouhov, T., & Muthén, B. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling*, *21*(4), 495–508. https://doi.org/10.1080/10705511.2014.919210

Atari, M., Graham, J., & Dehghani, M. (2020). Foundations of morality in Iran. *Evolution and Human Behavior*, *41*(5), 367–384. https://doi.org/10.1016/j.evolhumbehav.2020.07.014

Atari, M., Lai, M. H., & Dehghani, M. (2020). Sex differences in moral judgements across 67 countries. *Proceedings of the Royal Society B*, *287*(1937), Article 20201201. https://doi.org/10.1098/rspb.2020.1201

Atari, M., Mostafazadeh Davani, A., & Dehghani, M. (2020). Body maps of moral concerns. *Psychological Science*, *31*(2), 160–169. https://doi.org/10.1177/0956797619895284

Atari, M., Reimer, N. K., Graham, J., Hoover, J., Kennedy, B., Davani, A. M., Karimi-Malekabadi, F., Birjandi, S., & Dehghani, M. (2022). Pathogens are linked to human moral systems across time and space. *Current Research in Ecological and Social Psychology*, *3*, Article 100060. https://doi.org/10.1016/j.cresp.2022.100060

Baumard, N., André, J. B., & Sperber, D. (2013). A mutualistic approach to morality: The evolution of fairness by partner choice. *Behavioral and Brain Sciences*, *36*(1), 59–78. https://doi.org/10.1017/S0140525X11002202

Bazerman, M., White, S., & Lowenstein, G. (1995). Perceptions of fairness in interpersonal and individual choice situations. *Current Directions in Psychological Science*, *4*(2), 39–43. https://doi.org/10.1111/1467-8721.ep10770996

Beer, A., & Watson, D. (2009). The Individual and Group Loyalty Scales (IGLS): Construction and preliminary validation. *Journal of Personality Assessment*, *91*(3), 277–287. https://doi.org/10.1080/00223890902794341

Benenson, J. F., Webb, C. E., & Wrangham, R. W. (2022). Self-protection as an adaptive female strategy. *Behavioral and Brain Sciences*, *45*, 1–86. https://doi.org/10.1017/S0140525X21002417

Bloom, P. (2012). Religion, morality, evolution. *Annual Review of Psychology*, *63*(1), 179–199. https://doi.org/10.1146/annurev-psych-120710-100334

Brandt, M. J., Sibley, C. G., & Osborne, D. (2019). What is central to political belief system networks? *Personality and Social Psychology Bulletin*, *45*(9), 1352–1364. https://doi.org/10.1177/0146167218824354

Bringmann, L. F., & Eronen, M. I. (2018). Don't blame the model: Reconsidering the network approach to psychopathology. *Psychological Review*, *125*(4), 606–615. https://doi.org/10.1037/rev0000108

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, *1*(3), 185–216. https://doi.org/10.1177/135910457000100301

Brosnan, S. F. (2013). Justice- and fairness-related behaviors in nonhuman primates. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(Suppl. 2), 10416–10423. https://doi.org/10.1073/pnas.1301194110

Burger, J., Isvoranu, A. M., Lunansky, G., Haslbeck, J. M. B., Epskamp, S., Hoekstra, R. H. A., Fried, E. I., Borsboom, D., & Blanken, T. F. (2022). Reporting standards for psychological network analyses in cross-sectional data. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000471

Camerer, C., & Thaler, R. (1995). Ultimatums, dictators, and manners. *The Journal of Economic Perspectives*, *9*(2), 209–219. https://doi.org/10.1257/jep.9.2.209

Chester, D. S., & Lasko, E. N. (2021). Construct validation of experimental manipulations in social psychology: Current practices and recommendations for the future. *Perspectives on Psychological Science*, *16*(2), 377–395. https://doi.org/10.1177/1745691620950684

Christensen, A. P., Golino, H., & Silvia, P. J. (2020). A psychometric network perspective on the validity and validation of personality trait questionnaires. *European Journal of Personality*, *34*(6), 1095–1108. https://doi.org/10.1002/per.2265

Christie, N. C., Hsu, E., Iskiwitch, C., Iyer, R., Graham, J., Schwartz, B., & Monterosso, J. R. (2019). The moral foundations of needle exchange attitudes. *Social Cognition*, *37*(3), 229–246. https://doi.org/10.1521/soco.2019.37.3.229

Chudek, M., & Henrich, J. (2011). Culture-gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences*, *15*(5), 218–226. https://doi.org/10.1016/j.tics.2011.03.003

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

Clifford, S., & Jerit, J. (2013). How words do the work of politics: Moral foundations theory and the debate over stem cell research. *The Journal of Politics*, *75*(3), 659–671. https://doi.org/10.1017/S0022381613000492

Clifton, A., & Webster, G. D. (2017). An introduction to social network analysis for personality and social psychologists. *Social Psychological & Personality Science*, *8*(4), 442–453. https://doi.org/10.1177/1948550617709114

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, *92*(1), 1–45. https://doi.org/10.1016/S0304-4076(98)00084-0

Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., & Cramer, A. O. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, *54*, 13–29. https://doi.org/10.1016/j.jrp.2014.07.003

Costello, T. H., Bowes, S. M., Stevens, S. T., Waldman, I. D., Tasimi, A., & Lilienfeld, S. O. (2022). Clarifying the structure and nature of left-wing authoritarianism. *Journal of Personality and Social Psychology*, *122*(1), 135–170. https://doi.org/10.1037/pspp0000341

Crone, D. L. (2022, April 30). *Conceptual issues with the moral foundation of Purity: The case of religion*. PsyArXiv. https://doi.org/10.31234/osf.io/3e8bv

Crone, D. L., Bode, S., Murawski, C., & Laham, S. M. (2018). The Socio-Moral Image Database (SMID): A novel stimulus set for the study of social, moral and affective processes. *PLOS ONE*, *13*(1), Article e0190954. https://doi.org/10.1371/journal.pone.0190954

Crone, D. L., & Laham, S. M. (2022). Clarifying measurement issues with the purity subscale of the moral foundations questionnaire in Christian and nonreligious participants. *Social Psychological & Personality Science*. Advance online publication. https://doi.org/10.1177/19485506221136664

Cruza-Guet, M.-C., Spokane, A. R., León-Andrade, C., & Borja, T. (2009). Diversity, hegemony, poverty, and the emergence of counseling psychology in Ecuador. In L. H. Gerstein, P. P. Heppner, S. Ægisdóttir, S.-M. A. Leung, & K. L. Norsworthy (Eds.), *International handbook of cross-cultural counseling: Cultural assumptions and practices worldwide* (pp. 393–401). Sage Publications, Inc. https://doi.org/10.4135/.n30

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

Dalbert, C. (1999). The world is more just for me than generally: About the personal belief in a just world scale's validity. *Social Justice Research*, *12*(2), 79–98. https://doi.org/10.1023/A:1022091609047

Davey, L. M., Bobocel, D. R., Hing, L. S. S., & Zanna, M. P. (1999). Preference for the Merit Principle Scale: An individual difference measure of distributive justice preferences. *Social Justice Research*, *12*(3), 223–240. https://doi.org/10.1023/A:1022148418210

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., Dooley, M. T., Hook, J. N., Choe, E., & McElroy, S. E. (2017). The purity/sanctity subscale of the Moral Foundations Questionnaire does

not work similarly for religious versus non-religious individuals. *Psychology of Religion and Spirituality*, *9*(1), 124–130. https://doi.org/10.1037/rel0000057

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113–126. https://doi.org/10.1037/0022-3514.44.1.113

Del Giudice, M. (2009). On the real magnitude of psychological sex differences. *Evolutionary Psychology*, *7*(2), 264–279. https://doi.org/10.1177/147470490900700209

Del Giudice, M. (2022). Measuring sex differences and similarities. In D. P. VanderLaan & W. I. Wong (Eds.), *Gender and sexuality development: Contemporary theory and research* (pp. 1–38). Springer. https://doi.org/10.1007/978-3-030-84273-4_1

Deutsch, M. (1975). Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *Journal of Social Issues*, *31*(3), 137–149. https://doi.org/10.1111/j.1540-4560.1975.tb01000.x

Doğruyol, B., Alper, S., & Yilmaz, O. (2019). The five-factor model of the moral foundations theory is stable across WEIRD and non-WEIRD cultures. *Personality and Individual Differences*, *151*, Article 109547. https://doi.org/10.1016/j.paid.2019.109547

Durkee, P. K., Lukaszewski, A. W., von Rueden, C. R., Gurven, M. D., Buss, D. M., & Tucker-Drob, E. M. (2020). Niche diversity predicts personality structure across 115 nations. *Psychological Science*, *33*(2), 285–298. https://doi.org/10.1177/09567976211031571

Eagly, A. H., & Revelle, W. (2022). Understanding the magnitude of psychological differences between women and men requires seeing the forest and the trees. *Perspectives on Psychological Science*, *17*(5), 1339–1358. https://doi.org/10.1177/17456916211046006

Ebert, T., Götz, F. M., Mewes, L., & Rentfrow, P. J. (2022). Spatial analysis for psychologists: How to use individual-level data for research at the geographically aggregated level. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000493

Enke, B. (2020). Moral values and voting. *Journal of Political Economy*, *128*(10), 3679–3729. https://doi.org/10.1086/708857

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Fischer, R., & Karl, J. A. (2019). A primer to (cross-cultural) multi-group invariance testing possibilities in R. *Frontiers in Psychology*, *10*, Article 1507. https://doi.org/10.3389/fpsyg.2019.01507

Fiske, A. P. (1990). Relativity within Moose ("Mossi") culture: Four incommensurable models for social relationships. *Ethos*, *18*(2), 180–204. https://doi.org/10.1525/eth.1990.18.2.02a00040

Fiske, A. P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review*, *99*(4), 689–723. https://doi.org/10.1037/0033-295X.99.4.689

Flake, J. K., & Fried, E. I. (2020). Measurement schmeasurement: Questionable measurement practices and how to avoid them. *Advances in Methods and Practices in Psychological Science*, *3*(4), 456–465. https://doi.org/10.1177/2515245920952393

Flake, J. K., Pek, J., & Hehman, E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological & Personality Science*, *8*(4), 370–378. https://doi.org/10.1177/1948550617693063

Flora, D. B. (2020). Your coefficient alpha is probably wrong, but which coefficient omega is right? A tutorial on using R to obtain better reliability estimates. *Advances in Methods and Practices in Psychological Science*, *3*(4), 484–501. https://doi.org/10.1177/2515245920951747

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432–441. https://doi.org/10.1093/biostatistics/kxm045

Frimer, J. A., Boghrati, R., Haidt, J., Graham, J., & Dehghani, M. (2019). *Moral foundations dictionary for linguistic analyses 2.0* [Unpublished manuscript].

Fukuyama, F. (2006). *The end of history and the last man*. Simon and Schuster.

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLOS ONE*, *12*(6), Article e0174035. https://doi.org/10.1371/journal.pone.0174035

Golino, H. F., Shi, D., Christensen, A. P., Garrido, L. E., Nieto, M. D., Sadana, R., Thiyagarajan, J. A., & Martinez-Molina, A. (2020). Investigating the performance of exploratory graph analysis and traditional techniques to identify the number of latent factors: A simulation and tutorial. *Psychological Methods*, *25*(3), 292–320. https://doi.org/10.1037/met0000255

Goudarzi, S., Pliskin, R., Jost, J. T., & Knowles, E. D. (2020). Economic system justification predicts muted emotional responses to inequality. *Nature Communications*, *11*(1), Article 383. https://doi.org/10.1038/s41467-019-14193-z

Graham, J. (2015). Explaining away differences in moral judgment: Comment on Gray and Keeney (2015). *Social Psychological & Personality Science*, *6*(8), 869–873. https://doi.org/10.1177/1948550615592242

Graham, J., Atari, M., Dehghani, M., & Haidt, J. (in press). Puritanism needs purity, and moral psychology needs pluralism. *Behavioral and Brain Sciences*.

Graham, J., & Haidt, J. (2010). Beyond beliefs: Religions bind individuals into moral communities. *Personality and Social Psychology Review*, *14*(1), 140–150. https://doi.org/10.1177/1088868309353415

Graham, J., & Haidt, J. (2012). Sacred values and evil adversaries: A moral foundations approach. In M. Mikulincer & P. R. Shaver (Eds.), *The social psychology of morality: Exploring the causes of good and evil* (pp. 11–31). American Psychological Association. https://doi.org/10.1037/13091-001

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & 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. (2018). Moral foundations theory: On the advantages of moral pluralism over moral monism. In K. Gray & J. Graham (Eds.), *The atlas of moral psychology: Mapping good and evil in the mind* (pp. 211–222). Guilford Press.

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 and Social Psychology*, *101*(2), 366–385. https://doi.org/10.1037/a0021847

Gray, K., DiMaggio, N., Schein, C., & Kachanoff, F. (2022). The problem of purity in moral psychology. *Personality and Social Psychology Review*. Advance online publication. https://doi.org/10.1177/10888683221124741

Gray, K., Schein, C., & Ward, A. F. (2014). The myth of harmless wrongs in moral cognition: Automatic dyadic completion from sin to suffering. *Journal of Experimental Psychology: General*, *143*(4), 1600–1615. https://doi.org/10.1037/a0036149

Greene, J. D. (2013). *Moral tribes: Emotion, reason, and the gap between us and them*. Penguin.

Gurven, M. D. (2018). Broadening horizons: Sample diversity and socioecological theory are essential to the future of psychological science. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(45), 11420–11427. https://doi.org/10.1073/pnas.1720433115

Hafer, C. L., & Rubel, A. N. (2015). The why and how of defending belief in a just world. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. 51, pp. 41–96). Academic Press. https://doi.org/10.1016/bs.aesp.2014.09.001

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*(4), 814–834. https://doi.org/10.1037/0033-295X.108.4.814

Haidt, J. (2007). The new synthesis in moral psychology. *Science*, *316*(5827), 998–1002. https://doi.org/10.1126/science.1137651

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Vintage.

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, *20*(1), 98–116. https://doi.org/10.1007/s11211-007-0034-z

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, *133*(4), 55–66. https://doi.org/10.1162/0011526042365555

Haidt, J., & Joseph, C. (2011). How Moral Foundations Theory succeeded in building on sand: A response to Suhler and Churchland. *Journal of Cognitive Neuroscience*, *23*(9), 2117–2122. https://doi.org/10.1162/jocn.2011.21638

Haidt, J., McCauley, C., & Rozin, P. (1994). Individual differences in sensitivity to disgust: A scale sampling seven domains of disgust elicitors. *Personality and Individual Differences*, *16*(5), 701–713. https://doi.org/10.1016/0191-8869(94)90212-7

Harper, C. A., & Rhodes, D. (2021). Reanalysing the factor structure of the moral foundations questionnaire. *British Journal of Social Psychology*, *60*(4), 1303–1329. https://doi.org/10.1111/bjso.12452

Hatemi, P. K., Crabtree, C., & Smith, K. B. (2019). Ideology justifies morality: Political beliefs predict moral foundations. *American Journal of Political Science*, *63*(4), 788–806. https://doi.org/10.1111/ajps.12448

Henrich, J. (2020). *The WEIRDest people in the world: How the west became psychologically peculiar and particularly prosperous*. Farrar, Straus & Giroux.

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

Ho, A. K., Sidanius, J., Kteily, N., Sheehy-Skeffington, J., Pratto, F., Henkel, K. E., Foels, R., & Stewart, A. L. (2015). The nature of social dominance orientation: Theorizing and measuring preferences for intergroup inequality using the new SDO_7 scale. *Journal of Personality and Social Psychology*, *109*(6), 1003–1028. https://doi.org/10.1037/pspi0000033

Homans, G. C. (1961). *Social behavior: Its elementary forms*. Harcourt, Brace, and World.

Hoover, J., Atari, M., Mostafazadeh Davani, A., Kennedy, B., Portillo-Wightman, G., Yeh, L., & Dehghani, M. (2021). Investigating the role of group-based morality in extreme behavioral expressions of prejudice. *Nature Communications*, *12*(1), Article 4585. https://doi.org/10.1038/s41467-021-24786-2

Hoover, J., Portillo-Wightman, G., Yeh, L., Havaldar, S., Davani, A. M., Lin, Y., Kennedy, B., Atari, M., Kamel, Z., Mendlen, M., Moreno, G., Park, C., Chang, T. E., Chin, J., Leong, C., Leung, J. Y., Mirinjian, A., & Dehghani, M. (2020). Moral Foundations Twitter Corpus: A collection of 35k tweets annotated for moral sentiment. *Social Psychological & Personality Science*, *11*(8), 1057–1071. https://doi.org/10.1177/1948550619876629

Huberty, C. J. (2005). Mahalanobis distance. In B. S. Everitt & D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science* (pp. 1110–1111). Wiley.

Hunter, J. D. (1991). *Culture wars: The struggle to define America*. Basic Books.

Iurino, K., & Saucier, G. (2020). Testing measurement invariance of the Moral Foundations Questionnaire across 27 countries. *Assessment*, *27*(2), 365–372. https://doi.org/10.1177/1073191118817916

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), Article 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

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

Jost, J. T., & Thompson, E. P. (2000). Group-based dominance and opposition to equality as independent predictors of self-esteem, ethnocentrism, and social policy attitudes among African Americans and European Americans. *Journal of Experimental Social Psychology*, *36*(3), 209–232. https://doi.org/10.1006/jesp.1999.1403

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Strauss, Giroux.

Kennedy, B., Atari, M., Mostafazadeh Davani, A., Hoover, J., Omrani, A., Graham, J., & Dehghani, M. (2021). Moral concerns are differentially observable in language. *Cognition*, *212*, Article 104696. https://doi.org/10.1016/j.cognition.2021.104696

Khoudary, A., Hanna, E., O'Neill, K., Iyengar, V., Clifford, S., Cabeza, R., De Brigard, F., & Sinnott-Armstrong, W. (2022). A functional neuroimaging investigation of Moral Foundations Theory. *Social Neuroscience*, *17*(6), 491–507. https://doi.org/10.1080/17470919.2022.2148737

Kivikangas, J. M., Fernández-Castilla, B., Järvelä, S., Ravaja, N., & Lönnqvist, J. E. (2021). Moral foundations and political orientation: Systematic review and meta-analysis. *Psychological Bulletin*, *147*(1), 55–94. https://doi.org/10.1037/bul0000308

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Alper, S., Aveyard, M., Axt, J. R., Babalola, M. T., Bahník, Š., Batra, R., Berkics, M., Bernstein, M. J., Berry, D. R., Bialobrzeska, O., Binan, E. D., Bocian, K., Brandt, M. J., Busching, R., Rédei, A. C., … Nosek, B. A. (2018). Many Labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science*, *1*(4), 443–490. https://doi.org/10.1177/2515245918810225

Koenig, H., Parkerson, G. R., Jr., & Meador, K. G. (1997). Religion index for psychiatric research. *The American Journal of Psychiatry*, *154*(6), 885–886. https://doi.org/10.1176/ajp.154.6.885b

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347–480). Rand McNally.

Koleva, S. P., Graham, J., Iyer, R., Ditto, P. H., & Haidt, J. (2012). Tracing the threads: How five moral concerns (especially Purity) help explain culture war attitudes. *Journal of Research in Personality*, *46*(2), 184–194. https://doi.org/10.1016/j.jrp.2012.01.006

Krosnick, J. A., & Fabrigar, L. R. (1997). Designing rating scales for effective measurement in surveys. In L. Lyberg, P. Biemer, M. Collins, E. De Leeuw, C. Dippo, N. Schwarz, & D. Trewin (Eds.), *Survey measurement and process quality* (pp. 141–164). Wiley. https://doi.org/10.1002/9781118490013.ch6

Lee, N., & Cadogan, J. W. (2013). Problems with formative and higher-order reflective variables. *Journal of Business Research*, *66*(2), 242–247. https://doi.org/10.1016/j.jbusres.2012.08.004

Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology*, *68*(1), 151–158. https://doi.org/10.1037/0022-3514.68.1.151

Litman, L., & Robinson, J. (2020). *Conducting online research on Amazon Mechanical Turk and beyond*. Sage Publications.

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, *49*(2), 433–442. https://doi.org/10.3758/s13428-016-0727-z

Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, *10*(1), 85–110. https://doi.org/10.1146/annurev-clinpsy-032813-153700

McAdams, D. P., Albaugh, M., Farber, E., Daniels, J., Logan, R. L., & Olson, B. (2008). Family metaphors and moral intuitions: How conservatives and

liberals narrate their lives. *Journal of Personality and Social Psychology*, *95*(4), 978–990. https://doi.org/10.1037/a0012650

McCurrie, C. H., Crone, D. L., Bigelow, F., & Laham, S. M. (2018). Moral and Affective Film Set (MAAFS): A normed moral video database. *PLOS ONE*, *13*(11), Article e0206604. https://doi.org/10.1371/journal.pone.0206604

Meade, A. W., & Bauer, D. J. (2007). Power and precision in confirmatory factor analytic tests of measurement invariance. *Structural Equation Modeling*, *14*(4), 611–635. https://doi.org/10.1080/10705510701575461

Medin, D., Bennis, W., & Chandler, M. (2010). Culture and the home-field disadvantage. *Perspectives on Psychological Science*, *5*(6), 708–713. https://doi.org/10.1177/1745691610388772

Meindl, P., Iyer, R., & Graham, J. (2019). Distributive justice beliefs are guided by whether people think the ultimate goal of society is well-being or power. *Basic and Applied Social Psychology*, *41*(6), 359–385. https://doi.org/10.1080/01973533.2019.1663524

Meyers, J. L., & Beretvas, S. N. (2006). The impact of inappropriate modeling of cross-classified data structures. *Multivariate Behavioral Research*, *41*(4), 473–497. https://doi.org/10.1207/s15327906mbr4104_3

Morin, A. J., Arens, A. K., & Marsh, H. W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling*, *23*(1), 116–139. https://doi.org/10.1080/10705511.2014.961800

Moshontz, H., Campbell, L., Ebersole, C. R., IJzerman, H., Urry, H. L., Forscher, P. S., Grahe, J. E., McCarthy, R. J., Musser, E. D., Antfolk, J., Castille, C. M., Evans, T. R., Fiedler, S., Flake, J. K., Forero, D. A., Janssen, S. M. J., Keene, J. R., Protzko, J., Aczel, B., … Chartier, C. R. (2018). The Psychological Science Accelerator: Advancing psychology through a distributed collaborative network. *Advances in Methods and Practices in Psychological Science*, *1*(4), 501–515. https://doi.org/10.1177/2515245918797607

Muthén, B., & Asparouhov, T. (2014). IRT studies of many groups: The alignment method. *Frontiers in Psychology*, *5*, Article 978. https://doi.org/10.3389/fpsyg.2014.00978

Muthukrishna, M., Bell, A. V., Henrich, J., Curtin, C. M., Gedranovich, A., McInerney, J., & Thue, B. (2020). Beyond Western, Educated, Industrial, Rich, and Democratic (WEIRD) psychology: Measuring and mapping scales of cultural and psychological distance. *Psychological Science*, *31*(6), 678–701. https://doi.org/10.1177/0956797620916782

Nejat, P., & Hatami, J. (2019). Psychometric properties of the Persian version of moral foundations questionnaire in three Iranian samples. *Social Cognition*, *8*(1), 107–124. https://doi.org/10.30473/SC.2019.40617.2204

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. (2020). Moral foundations theory and the psychology of charitable giving. *European Journal of Personality*, *34*(3), 431–447. https://doi.org/10.1002/per.2256

Norenzayan, A. (2016). Theodiversity. *Annual Review of Psychology*, *67*(1), 465–488. https://doi.org/10.1146/annurev-psych-122414-033426

Olatunji, B. O., Williams, N. L., Tolin, D. F., Abramowitz, J. S., Sawchuk, C. N., Lohr, J. M., & Elwood, L. S. (2007). The Disgust Scale: Item analysis, factor structure, and suggestions for refinement. *Psychological Assessment*, *19*(3), 281–297. https://doi.org/10.1037/1040-3590.19.3.281

Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. A. (2014). The role of analytic thinking in moral judgements and values. *Thinking & Reasoning*, *20*(2), 188–214. https://doi.org/10.1080/13546783.2013.865000

Petersen, M. B., Sznycer, D., Sell, A., Cosmides, L., & Tooby, J. (2013). The ancestral logic of politics: Upper-body strength regulates men's assertion of self-interest over economic redistribution. *Psychological Science*, *24*(7), 1098–1103. https://doi.org/10.1177/0956797612466415

Pons, P., & Latapy, M. (2006). Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications*, *10*(2), 191–218. https://doi.org/10.7155/jgaa.00124

Purzycki, B. G., Pisor, A. C., Apicella, C., Atkinson, Q., Cohen, E., Henrich, J., McElreath, R., McNamara, R. A., Norenzayan, A., Willard, A. K., & Xygalatas, D. (2018). The cognitive and cultural foundations of moral behavior. *Evolution and Human Behavior*, *39*(5), 490–501. https://doi.org/10.1016/j.evolhumbehav.2018.04.004

Rai, T. S. (2018). Relationship regulation theory. In K. Gray & J. Graham (Eds.), *Atlas of moral psychology* (pp. 231–241). Guilford Press.

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review*, *118*(1), 57–75. https://doi.org/10.1037/a0021867

Reimer, N. K., Atari, M., Karimi-Malekabadi, F., Trager, J., Kennedy, B., Graham, J., & Dehghani, M. (2022). Moral values predict county-level COVID-19 vaccination rates in the United States. *American Psychologist*, *77*(6), 743–759. https://doi.org/10.1037/amp0001020

Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika*, *74*(1), 145–154. https://doi.org/10.1007/s11336-008-9102-z

Rozin, P., Haidt, J., & McCauley, C. R. (2000). Disgust. In M. Lewis & J. Haviland (Eds.), *Handbook of emotions* (2nd ed., pp. 637–653). Guilford Press.

Sandel, M. J. (2020). *The tyranny of merit: What's become of the common good?* Allen Lane.

Saroglou, V., & Cohen, A. B. (2013). Cultural and cross-cultural psychology of religion. In R. F. Paloutzian & C. L. Park (Eds.), *Handbook of the psychology of religion and spirituality* (pp. 330–354). Guilford Press.

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. *Advances in Experimental Social Psychology*, *25*, 1–65. https://doi.org/10.1016/S0065-2601(08)60281-6

Schwartz, S. H., & Rubel, T. (2005). Sex differences in value priorities: Cross-cultural and multimethod studies. *Journal of Personality and Social Psychology*, *89*(6), 1010–1028. https://doi.org/10.1037/0022-3514.89.6.1010

Shweder, R. A., & Haidt, J. (1993). The future of moral psychology: Truth, intuition, and the pluralist way. *Psychological Science*, *4*(6), 360–365. https://doi.org/10.1111/j.1467-9280.1993.tb00582.x

Shweder, R. A., Mahapatra, M., & Miller, J. (1987). Culture and moral development. In J. Kagan & S. Lamb (Eds.), *The emergence of morality in young children* (pp. 1–83). University of Chicago Press.

Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). The "big three" of morality (autonomy, community, and divinity), and the "big three" explanations of suffering. In A. Brandt & P. Rozin (Eds.), *Morality and health* (pp. 119–169). Routledge.

Skurka, C., Winett, L. B., Jarman-Miller, H., & Niederdeppe, J. (2020). All things being equal: Distinguishing proportionality and equity in moral reasoning. *Social Psychological & Personality Science*, *11*(3), 374–387. https://doi.org/10.1177/1948550619862261

Smaldino, P. E., Lukaszewski, A., von Rueden, C., & Gurven, M. (2019). Niche diversity can explain cross-cultural differences in personality structure. *Nature Human Behaviour*, *3*(12), 1276–1283. https://doi.org/10.1038/s41562-019-0730-3

Snarey, J. R., Reimer, J., & Kohlberg, L. (1985). Development of social-moral reasoning among Kibbutz adolescents: A longitudinal cross-cultural study. *Developmental Psychology*, *21*(1), 3–17. https://doi.org/10.1037/0012-1649.21.1.3

Starmans, C., Sheskin, M., & Bloom, P. (2017). Why people prefer unequal societies. *Nature Human Behaviour*, *1*(4), 1–7. https://doi.org/10.1038/s41562-017-0082

Stich, S. (2006). Is morality an elegant machine or a kludge? *Journal of Cognition and Culture*, *6*(1–2), 181–189. https://doi.org/10.1163/156853706776931349

Stone, W. F. (1980). The myth of left-wing authoritarianism. *Political Psychology*, *2*(3/4), 3–19. https://doi.org/10.2307/3790998

Sunar, D., Cesur, S., Piyale, Z. E., Tepe, B., Biten, A. F., Hill, C. T., & Koç, Y. (2021). People respond with different moral emotions to violations in different relational models: A cross-cultural comparison. *Emotion*, *21*(4), 693–706. https://doi.org/10.1037/emo0000736

Thalmayer, A. G., Toscanelli, C., & Arnett, J. J. (2021). The neglected 95% revisited: Is American psychology becoming less American? *American Psychologist*, *76*(1), 116–129. https://doi.org/10.1037/amp0000622

Triandis, H. C. (1995). *Individualism and collectivism*. Westview.

Triandis, H. C., & Gelfand, M. J. (1998). Converging measurement of horizontal and vertical individualism and collectivism. *Journal of Personality and Social Psychology*, *74*(1), 118–128. https://doi.org/10.1037/0022-3514.74.1.118

Turner-Zwinkels, F. M., Johnson, B. B., Sibley, C. G., & Brandt, M. J. (2021). Conservatives' moral foundations are more densely connected than liberals' moral foundations. *Personality and Social Psychology Bulletin*, *47*(2), 167–184. https://doi.org/10.1177/0146167220916070

von Rueden, C., Alami, S., Kaplan, H., & Gurven, M. (2018). Sex differences in political leadership in an egalitarian society. *Evolution and Human Behavior*, *39*(4), 402–411. https://doi.org/10.1016/j.evolhumbehav.2018.03.005

Ward, M. D., & Gleditsch, K. S. (2008). *Spatial regression models*. Sage Publications. https://doi.org/10.4135/9781412985888

White, C. J. M., Kelly, J. M., Shariff, A. F., & Norenzayan, A. (2019). Supernatural norm enforcement: Thinking about karma and God reduces selfishness among believers. *Journal of Experimental Social Psychology*, *84*, Article 103797. https://doi.org/10.1016/j.jesp.2019.03.008

White, C. J. M., Muthukrishna, M., & Norenzayan, A. (2021). Cultural similarity among coreligionists within and between countries. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(37), Article e2109650118. https://doi.org/10.1073/pnas.2109650118

Whitehead, H. (2000). *Food rules: Hunting, sharing, and tabooing game in Papua New Guinea*. University of Michigan Press. https://doi.org/10.3998/mpub.10823

Willard, A. K., Baimel, A., Turpin, H., Jong, J., & Whitehouse, H. (2020). Rewarding the good and punishing the bad: The role of karma and afterlife beliefs in shaping moral norms. *Evolution and Human Behavior*, *41*(5), 385–396. https://doi.org/10.1016/j.evolhumbehav.2020.07.001

Yalçındağ, B., Özkan, T., Cesur, S., Yilmaz, O., Tepe, B., Piyale, Z. E., Biten, A. F., & Sunar, D. (2019). An investigation of moral foundations theory in Turkey using different measures. *Current Psychology*, *38*(2), 440–457. https://doi.org/10.1007/s12144-017-9618-4

Yan, V. X., Oyserman, D., Kiper, G., & Atari, M. (2023). Difficulty-as-improvement: The courage to keep going in the face of life's difficulties. *Personality and Social Psychology Bulletin*. Advance online publication. https://doi.org/10.1177/01461672231153680

Yi, D., & Tsang, J. A. (2020). The relationship between individual differences in religion, religious primes, and the moral foundations. *Archiv für Religionspsychologie*, *42*(2), 161–193. https://doi.org/10.1177/0084672420909459

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

Zhang, Z. (2020). The impact of social relationships on children's distributive justice. *Developmental Psychology*, *56*(1), 103–116. https://doi.org/10.1037/dev0000855

Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale's indicators: A comparison of estimators for ωh. *Applied Psychological Measurement*, *30*(2), 121–144. https://doi.org/10.1177/0146621605278814

Appendix

Moral Foundations Questionnaire–2 (MFQ-2)

For each of the statements below, please indicate how well each statement describes you or your opinions. Response options: *Does not describe me at all* (1); *slightly describes me* (2); *moderately describes me* (3); *describes me fairly well* (4); and *describes me extremely well* (5).

1. Caring for people who have suffered is an important virtue.
2. The world would be a better place if everyone made the same amount of money.
3. I think people who are more hardworking should end up with more money.
4. I think children should be taught to be loyal to their country.
5. I think it is important for societies to cherish their traditional values.
6. I think the human body should be treated like a temple, housing something sacred within.
7. I believe that compassion for those who are suffering is one of the most crucial virtues.
8. Our society would have fewer problems if people had the same income.
9. I think people should be rewarded in proportion to what they contribute.
10. It upsets me when people have no loyalty to their country.
11. I feel that most traditions serve a valuable function in keeping society orderly.
12. I believe chastity is an important virtue.
13. We should all care for people who are in emotional pain.
14. I believe that everyone should be given the same quantity of resources in life.
15. The effort a worker puts into a job ought to be reflected in the size of a raise they receive.

(*Appendix continues*)

16. Everyone should love their own community.
17. I think obedience to parents is an important virtue.
18. It upsets me when people use foul language like it is nothing.
19. I am empathetic toward those people who have suffered in their lives.
20. I believe it would be ideal if everyone in society wound up with roughly the same amount of money.
21. It makes me happy when people are recognized on their merits.
22. Everyone should defend their country, if called upon.
23. We all need to learn from our elders.
24. If I found out that an acquaintance had an unusual but harmless sexual fetish I would feel uneasy about them.
25. Everyone should try to comfort people who are going through something hard.
26. When people work together toward a common goal, they should share the rewards equally, even if some worked harder on it.
27. In a fair society, those who work hard should live with higher standards of living.
28. Everyone should feel proud when a person in their community wins in an international competition.
29. I believe that one of the most important values to teach children is to have respect for authority.
30. People should try to use natural medicines rather than chemically identical human-made ones.
31. It pains me when I see someone ignoring the needs of another human being.
32. I get upset when some people have a lot more money than others in my country.
33. I feel good when I see cheaters get caught and punished.
34. I believe the strength of a sports team comes from the loyalty of its members to each other.
35. I think having a strong leader is good for society.
36. I admire people who keep their virginity until marriage.

Scoring: Average each of the following items to get six scores corresponding with the six foundations.

Care = 1, 7, 13, 19, 25, and 31
Equality = 2, 8, 14, 20, 26, and 32
Proportionality = 3, 9, 15, 21, 27, and 33
Loyalty = 4, 10, 16, 22, 28, and 34
Authority = 5, 11, 17, 23, 29, and 35
Purity = 6, 12, 18, 24, 30, and 36

Received June 18, 2022
Revision received May 7, 2023
Accepted May 10, 2023 ■