Full Length Article

Mapping morality with a compass: Testing the theory of 'morality-as-cooperation' with a new questionnaire ☆

Oliver Scott Curry [a,*], Matthew Jones Chesters [b], Caspar J. Van Lissa [c]

[a] Institute of Cognitive and Evolutionary Anthropology, University of Oxford, United Kingdom
[b] School of Psychology, University of East London, United Kingdom
[c] Department of Methodology & Statistics, Utrecht University, Netherlands

ABSTRACT

Morality-as-Cooperation (MAC) is the theory that morality is a collection of biological and cultural solutions to the problems of cooperation recurrent in human social life. MAC uses game theory to identify distinct types of cooperation, and predicts that each will be considered morally relevant, and each will give rise to a distinct moral domain. Here we test MAC's predictions by developing a new self-report measure of morality, the Morality-as-Cooperation Questionnaire (MAC-Q), and comparing its psychometric properties to those of the Moral Foundations Questionnaire (MFQ). Over four studies, the results support the MAC-Q's seven-factor model of morality, but not the MFQ's five-factor model. Thus MAC emerges as the best available compass with which to explore the moral landscape.
© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

What is morality? What explains its content and structure? And how is it best measured? In recent years, the study of morality has become the focus of a thriving interdisciplinary endeavour, encompassing research not only in psychology, but also in evolutionary theory, genetics, biology, animal behaviour, anthropology, neuroscience and economics (Haidt, 2007; Shackelford & Hansen, 2016; Sinnott-Armstrong, 2007). A common view in this body of work is that the function of morality is to promote cooperation (Curry, 2016; Greene, 2015:40; Haidt & Kesebir, 2010:800; Rai & Fiske, 2011:59; Sterelny & Fraser, 2016:1; Tomasello & Vaish, 2013:231).[1]

However, previous cooperative accounts of morality have not made full use of the mathematical analysis of cooperation – the theory of nonzerosum games – to provide a systematic taxonomy of cooperation. They have instead tended to focus on a relatively narrow range of cooperative behaviours (typically kin altruism and reciprocal altruism), and omitted others (for example, coordination and conflict resolution) (Table 4 in Curry, 2016). Thus, previous accounts have attempted to explain morality from an unnecessarily restricted base, and missed the opportunity to furnish a broader, more general theory of morality.

The present paper has two goals. First, we use nonzerosum game theory to provide the rigorous, systematic foundation that the cooperative approach to morality has previously lacked. We show how this rich, principled explanatory framework – which we call 'Morality-as-Cooperation' (MAC; Curry, 2016; Curry, Mullins, & Whitehouse, 2019) – incorporates more types of cooperation, and thus explains more types of morality, than previous approaches. The current version of the theory incorporates seven well-established types of cooperation: (1) the allocation of resources to kin (Hamilton, 1963); (2) coordination to mutual advantage (Lewis, 1969); (3) social exchange (Trivers, 1971); and

---

[1] "Moral systems are interlocking sets of values, virtues, norms, practices, identities, institutions, technologies, and evolved psychological mechanisms that work together to suppress or regulate selfishness and make *cooperative* social life possible." (Haidt & Kesebir, 2010). "[M]orality functions to facilitate the generation and maintenance of long-term social-*cooperative* relationships" (Rai & Fiske, 2011). "Human morality arose evolutionarily as a set of skills and motives for *cooperating with others*" (Tomasello & Vaish, 2013). "[T]he core function of morality is to promote and sustain *cooperation*" (Greene, 2015). "[M]oral facts are facts about *cooperation*, and the conditions and practices that support or undermine it" (Sterelny & Fraser, 2016). Emphasis added.

conflict resolution through contests featuring displays of (4) hawk-ish and (5) dove-ish traits (Maynard Smith & Price, 1973); (6) division (Skyrms, 1996); and (7) possession (Gintis, 2007).

Second, we test MAC's prediction that each of these types of cooperation will be considered morally relevant, and each will give rise to a distinct moral domain, by developing a new self-report measure of moral values – with facets dedicated to (1) family values, (2) group loyalty, (3) reciprocity, (4) bravery, (5) respect, (6) fairness and (7) property rights – and examine its psychometric properties.

2. How cooperation explains morality

The theory of Morality-as-Cooperation (MAC) argues that morality consists of a collection of biological and cultural solutions to the problems of cooperation recurrent in human social life (Curry, 2016). Below we review the general argument, before looking at how specific types of cooperation explain corresponding types of morality.

Life begins when molecules start making copies of themselves. These 'replicators' are 'selfish' in the technical sense that they promote their own replication (Dawkins, 1976/2006). They can promote their replication at the *expense* of other replicators. These *competitive* interactions have a winner and a loser; one's gain is another's loss; they are zerosum games (Maynard Smith, 1982; Von Neumann & Morgenstern, 1944). But replicators can also replicate in *concert* with other replicators (Dawkins, 1998). These *cooperative* interactions can have two winners; they are win-win situations; they are nonzerosum games. Natural selection can favour genes for cooperation – that is, genes for evolutionarily-stable phenotypic strategies designed to achieve superior equilibria in nonzerosum interactions – and has done throughout the history of life. Natural selection for genes that employ cooperative strategies has driven several 'major transitions' in the evolution of life on Earth, including the formation of cells, chromosomes and multicellular organisms (Maynard Smith & Szathmáry, 1995). Natural selection has also favoured genes for cooperation between individuals, in a wide variety of species (Dugatkin, 1997), including humans. Humans descend from a long line of social primates; they have spent 50 million years living in social groups (Shultz, Opie, & Atkinson, 2011), and two million years making a living as intensely collaborative hunter-gatherers (Tooby & DeVore, 1987). This has equipped humans with a range of biological – including psychological – adaptations for cooperation. These adaptations can be seen as natural selection's 'attempts' to solve the problems of cooperation. More recently, improvisational intelligence and cultural transmission (Boyd, Richerson, & Henrich, 2011; Pinker, 2010) have made it possible for humans to attempt to improve upon natural selection's solutions by inventing evolutionarily-novel solutions – 'tools and rules' – for further bolstering cooperation (Binmore, 1994a, 1994b; Hammerstein, 2003; Nagel, 1991; Popper, 1945). Together, these biological and cultural mechanisms provide the motivation for social, cooperative and altruistic behaviour; and they provide the criteria by which individuals evaluate the behaviour of others. According to MAC, it is precisely these solutions to problems of cooperation – this collection of instincts, intuitions, inventions and institutions – that constitute human morality (Curry, 2005, 2016).[2]

Which problems of cooperation do humans face? And how are they solved? Evolutionary biology and game theory tell us that there is not just one problem of cooperation but many, with many different functionally, and perhaps phenotypically, distinct solutions (Lehmann & Keller, 2006; Nunn & Lewis, 2001; Robinson & Goforth, 2005; Sachs, Mueller, Wilcox, & Bull, 2004). Our review of this literature suggests that there are (at least) seven well-established types of cooperation: (1) the allocation of resources to kin; (2) coordination to mutual advantage; (3) social exchange; and conflict resolution through contests featuring (4) hawkish displays of dominance and (5) dove-ish displays of submission; (6) division of disputed resources; and (7) recognition of possession. We briefly review each of these below, and we consider how each type of cooperation provides an explanation for a corresponding type of morality (Table 1).

2.1. Allocation of resources to kin (Family Values)

Genes that benefit replicas of themselves that reside in other individuals – that is, genetic relatives – will be favoured by natural selection if the cost of helping is outweighed by the benefit to the recipient gene(s) (Dawkins, 1979; Hamilton, 1963). So, evolutionary theory leads us to expect that under some conditions organisms will possess adaptations for detecting and delivering benefits (or avoiding doing harm) to kin. This theory of kin selection explains many instances of altruism, in many species (Gardner & West, 2014), including humans (Kurland & Gaulin, 2005; Lieberman, Tooby, & Cosmides, 2007). MAC predicts that because strategies for kin altruism realise a mutual benefit, they will be regarded as morally good. This theory can explain why caring for offspring (Edel & Edel, 1959/1968; an 'ethic of care'; Gilligan, 1982), helping family members (Fukuyama, 1996; Wong, 1984), and avoiding inbreeding (Lieberman, Tooby, & Cosmides, 2003; Westermarck, 1994) have been widely regarded as important components of morality.

2.2. Coordination to mutual advantage (Group Loyalty)

In game theory, situations in which individuals are uncertain about how to behave in order to bring about a mutual benefit are modelled as coordination problems (Lewis, 1969). Humans and other animals use a variety of strategies – such as focal points, traditions, leadership, signalling, badges of membership, and 'theory of mind' – to solve these problems (Alvard, 2001; Boos, Kolbe, Kappeler, & Ellwart, 2011; Curry & Jones Chesters, 2012; McElreath, Boyd, & Richerson, 2003), and form stable coalitions and alliances (Balliet, Wu, & De Dreu, 2014; Bissonnette et al., 2015; Harcourt & de Waal, 1992). MAC predicts that because solutions to coordination problems realise mutual benefits, they will be regarded as morally good. This theory can explain why participating in collaborative endeavours (Royce, 1908), favouring your own group (Bernhard, Fischbacher, & Fehr, 2006; Gert, 2013), and adopting local conventions (Gibbard, 1990a, 1990b) have been widely regarded as important components of morality.

2.3. Social exchange (Reciprocity)

In game theory, social dilemmas – prisoners dilemmas, public goods games, tragedies of the commons – arise when the fruits of cooperation are vulnerable to 'free riders', who accept the benefit of cooperation, without paying the cost (Ostrom & Walker, 2002). This problem can be overcome by a strategy of 'conditional cooperation' or 'reciprocal altruism', such as tit-for-tat (Axelrod, 1984; Trivers, 1971). Evidence for conditional cooperation has been found in numerous animal species (Carter, 2014), including humans (Cosmides & Tooby, 2005; Henrich et al., 2005; Jaeggi & Gurven, 2013). MAC predicts that because solutions to social dilemmas realise mutual benefits, they will be regarded as morally

---

[2] To be clear, there are many ways of promoting genetic survival and reproduction; some involve interpersonal cooperation, some do not. MAC hypothesises that it is only the (un)cooperative strategies that are considered (im)moral. Other non-cooperative ways of promoting survival and reproduction – such as strategies for choosing habitats or avoiding predators – are not.

Table 1
Overview of morality-as-cooperation.

	Label	Problem/Opportunity	Solution	Virtues	Vices	Epithet
1	Family	Kin selection	Kin Altruism	Duty of care, special obligations to kin	Incest, neglect	*Blood is thicker than water*
2	Group	Coordination	Mutualism	Loyalty, unity, solidarity, conformity	Betrayal, treason	*United we stand, divided we fall*
3	Reciprocity	Social Dilemma	Reciprocal Altruism	Reciprocity, trustworthiness, forgiveness	Cheating, ingratitude	*One good turn deserves another*
4	Heroism	Conflict Resolution (Contest)	Hawkish Displays	Bravery, fortitude, largesse	Cowardice, miserliness	*With great power comes great responsibility*
5	Deference	Conflict Resolution (Contest)	Dove-ish Displays	Respect, obedience, humility	Disrespect, hubris	*Blessed are the meek*
6	Fairness	Conflict Resolution (Bargaining)	Division	Fairness, impartiality, equality	Unfairness, favouritism	*Let's meet in the middle*
7	Property	Conflict Resolution (Possession)	Ownership	Respect for property, property rights	Theft, trespass	*Possession in nine-tenths of the law*

good. This theory can explain why reciprocity in general (Chilton & Neusner, 2009; Confucius, 1994), as well as its various subcomponents – trust (Baier, 1995), patience (Curry, Price, & Price, 2008), gratitude (Emmons, 2004), guilt (Gibbard, 1990b), apology (Ohtsubo & Watanabe, 2009), and forgiveness (Downie, 1965; Godfray, 1992; Richards, 1988) – have been widely regarded as important components of morality.

2.4. Contests between Hawks (Heroism) & 2.5 Doves (Deference)

Conflict over resources – food, territory, and mates (Huntingford & Turner, 1987) – presents organisms with an opportunity to cooperate by competing in less mutually-destructive ways (Maynard Smith & Price, 1973). There are three ways of achieving this: contests (featuring the display of hawkish and dove-ish traits), division, and possession.

Game theory has shown that conflicts can be settled through 'contests', in which individuals display reliable indicators of their 'fighting ability', and the weaker 'contestant' defers to the stronger (Gintis, Smith, & Bowles, 2001; Maynard Smith & Price, 1973). Such contests are widespread in nature (Hardy & Briffa, 2013; Riechert, 1998), and often form the basis of dominance hierarchies where resources are allocated by 'rank' (Preuschoft & van Schaik, 2000). Humans have a similar repertoire of status-related behaviours (Fiddick, Cummins, Janicki, Lee, & Erlich, 2013; Mazur, 2005; Sell, Tooby, & Cosmides, 2009), and culturally elaborated hierarchies (Boone, 1992; Rubin, 2000). MAC predicts that because hawkish displays of dominance, and dove-ish displays of submission, together realise mutual benefits, they will be regarded as morally good. This theory can explain why these two apparently contradictory sets of traits (Berlin, 1997) – the 'heroic virtues' of fortitude, bravery, skill, and wit, and the 'monkish virtues' of humility, deference, obedience, and respect – have been widely regarded as important components of morality (Curry, 2007; MacIntyre, 1981a, 1981b).

2.6. Division (Fairness)

When the contested resource is divisible, game theory models the situation as a 'bargaining problem' (Nash, 1950). Here, one solution is to divide the resource in proportion to the relative (bargaining) power of the protagonists (Skyrms, 1996). In the case of equally powerful individuals, this results in equal shares (Maynard Smith, 1982). Evidence for a 'sense of fairness' comes from non-human primates' adverse reactions to unequal treatment in economic games (Brosnan, 2013; Brosnan & de Waal, 2014). With regard to humans, rules such as "I cut, you choose", "meet in the middle", "split the difference", and "take turns", are ancient and widespread means of resolving disputes (Brams & Taylor, 1996). And 'equal shares' is a spontaneous and cross-culturally prevalent decision rule in economic games (Henrich et al., 2005) and similar situations (Messick, 1993). MAC predicts that because dividing resources avoids a costly fight, and therefore realises a mutual benefit, it will be regarded as morally good. This theory

can explain why fairness (Rawls, 1958) and willingness to compromise (Pennock & Chapman, 1979) have been widely regarded as important components of morality.

2.7. Possession (Property Rights)

Finally, game theory shows that conflicts over resources can be resolved by deference to prior possession (Gintis, 2007; Hare, Reeve, & Blossey, 2016; Maynard Smith, 1982). The recognition of prior possession is widespread in nature (Sherratt & Mesterton-Gibbons, 2015; Strassmann & Queller, 2014). Humans also defer to prior possession in vignette studies (DeScioli & Karpoff, 2015; Friedman & Neary, 2008), experimental games (the 'endowment effect'; Kahneman & Tversky, 1979), the law (Rose, 1985), and international relations (Johnson & Toft, 2014). Private property, in some form or other, appears to be a cross-cultural universal (Herskovits, 1952). MAC predicts that because deferring to prior possession avoids a costly fight, and therefore realises a mutual benefit, it will be regarded as morally good. This theory can explain why the right to own property and the prohibition of theft (Becker, 1977; Locke, 2000; Pennock & Chapman, 1980) have been widely regarded as an important components of morality.

3. Summary and predictions

Morality-as-Cooperation (MAC) is the theory that morality consists of a collection of biological and cultural solutions to the problems of cooperation recurrent in human social life. MAC draws upon the mathematics of cooperation to identify and distinguish between different types of cooperation, and thereby explain different facets of morality. The present review has identified seven types of cooperation, and hence seven candidate moral domains: obligations to family, group loyalty, reciprocity, bravery, respect, fairness, and property rights. Thus MAC can explain why specific forms of cooperative behaviour – helping kin, helping one's group, reciprocating costs and benefits, displaying 'hawkish' and dove-ish traits, dividing disputed resources, and respecting prior possession – are regarded as morally good, and why the corresponding forms of uncooperative behaviour – neglecting kin, betraying one's group, free-riding, cowardice, disrespect, unfairness and theft – are regarded as morally bad.

Starting from these first principles, MAC makes the following predictions about morality. First, with regard to content, MAC predicts that people will regard each type of cooperation as morally relevant; that is, as falling within the moral domain. Second, with regard to structure, MAC predicts that because the incidence and value of these different types of cooperation vary independently in social life (and are perhaps subserved by different psychological mechanisms) the strength of endorsement of each of the corresponding types of morality will vary independently too. In other words, each of these seven types of cooperation will give rise to a distinct moral domain. Accordingly, the theory predicts that moral values will exhibit a multifactorial structure, varying on these seven dimensions. Moreover, as a corollary of this prediction

regarding structure, MAC predicts that behaviour *not* tied to a specific type of cooperation will *not* constitute a distinct moral domain. These predictions about the content and structure of morality distinguish MAC from previous evolutionary and cooperative theories of morality.

3.0.1. Moral Foundations Theory

The most widely-used, and thus far most extensive, attempt to map the moral domain is Moral Foundations Theory (MFT; Haidt & Graham, 2007) operationalised in the Moral Foundations Questionnaire (MFQ; Graham et al., 2011). Like MAC, MFT takes a cooperative approach to morality, and maintains that there are many moral domains. But, unlike MAC, MFT does not derive its domains from any underlying theory of cooperation (Haidt & Joseph, 2011), and proposes only five: Care, Fairness, Ingroup, Authority and Purity.[3] Like MAC, MFT includes domains dedicated to group loyalty (Ingroup), deference (Authority) and fairness (Fairness). But unlike MAC, MFT does not include domains dedicated to family, reciprocity, heroism, or property. MFT has no foundation dedicated to kin altruism; the MFQ does have two items pertaining to kin, but they appear under Fairness and Ingroup. Nor has MFT any foundation dedicated to reciprocal altruism: MFT places reciprocity (a solution to iterated prisoners' dilemmas) and fairness (a solution to bargaining problems) under the same heading, and the MFQ has no items pertaining to reciprocity. MFT has no foundations, and the MFQ has no items, dedicated to hawkish displays of dominance, such as bravery. And the only mention of property occurs in an item about inheritance under the foundation of Fairness.

MFT also includes domains – Care and Purity – that are not related to a specific type of cooperation, and that MAC therefore predicts will not constitute coherent domains.

MAC predicts that moral psychology will be sensitive to the benefits (care, altruism) and costs (harms) of social interaction — for what is cooperation but a particular configuration of benefits and costs? But, as we have seen, MAC suggests that there are different types of benefits and costs — with different causes and consequences. For example, some 'harms', such as murder, are considered morally bad because they violate one or more cooperative principles (they break implicit social contracts against the use of force, and constitute an escalation of conflict, as opposed to its peaceful resolution). Other 'harms', such as punishment or self-defence, are considered morally good because they promote cooperation. This perspective suggests that it is a mistake to attempt to analyse benefits and costs in isolation, outside of their cooperative context, by placing them in a separate, generic domain dedicated to care or harm.

'Purity', meanwhile, has been described as the avoidance of "people with diseases, parasites [and] waste products" (Haidt & Joseph, 2004). It has no explicated connection to cooperation; on the contrary, it is regarded as an "odd corner" of morality precisely because it is not "concerned with how we treat other people" (Haidt & Joseph, 2004). By contrast, MAC suggests that the problem of avoiding pathogens (and other disgust-eliciting stimuli) is not a moral problem *per se*; instead, 'pure' or 'impure' behaviour is moralised only when it provides benefits, or imposes costs on, others –

for example, by putting their health at risk. So, avoiding rotten fruit on a tree is not a moral issue, but coughing in public without covering your mouth is. And, because there are many different ways in which disgusting behaviour might influence others – the problem of avoiding incest is not the same as the problem of avoiding people with poor personal hygiene – MAC suggests that it is a mistake to single out 'purity' as a separate, generic domain.

3.0.2. Relational Models Theory

Similarly, like MAC, Fiske's Relational Models Theory (RMT) takes a cooperative approach to morality, and maintains that there are many moral domains. But, unlike MAC, RMT does not derive its domains from any underlying theory of cooperation, and proposes only four: Unity, Hierarchy, Equality and Proportionality (Fiske & Rai, 2014; Rai & Fiske, 2011).[4] Unlike MAC, RMT's domain of Unity does not distinguish between family and group; Hierarchy does not distinguish between hawkish heroism and dove-ish deference; and Equality and Proportionality do not distinguish between reciprocity and fairness. Interestingly, like MAC, and unlike MFT, RMT argues that there are no distinct domains dedicated to 'harm' or 'purity'.[5]

3.0.3. Theory of Dyadic Morality

Unlike MAC (and MFT and RMT), Gray's Theory of Dyadic Morality (TDM) (Schein & Gray, 2018) does not take a cooperative approach to morality, but instead argues that the function of moral rules is to minimise harm to others (and is therefore a form of utilitarianism). TDM recognises that there may be different "genres" of harm that correspond to MFT's domains, but argues that all moral violations are processed by general-purpose psychological mechanisms, as opposed to distinct special-purpose mechanisms. Like MAC, and RMT, TDM does not accept MFT's claim that 'purity' is a distinct domain of morality – indeed, TDM has marshalled considerable evidence to suggest that 'impure' or disgusting acts are merely a particular form of harmful behaviour (Gray, Schein, & Ward, 2014).

3.0.4. Side-Taking Theory of Morality

Finally, like MAC, DeScioli and Kurzban's 'side-taking' theory of morality (STTM) agrees that cooperation explains moral *behaviour*: "evolutionary theories of morality [that] focus on understanding cooperation…do an excellent job of explaining why humans…care for offspring, cooperate in groups, trade favors, communicate honestly, and respect property" (DeScioli, 2016: 23). However, whereas MAC would argue that these cooperative theories also explain why people make and express moral

---

[3] Care/Harm is said to relate to "virtues such as kindness and compassion, and also in corresponding vices such as cruelty and aggression". Fairness/Reciprocity relates to the virtues of "fairness and justice", "individual rights and equality". Ingroup/Loyalty relates to "virtues such as loyalty, patriotism, and heroism" and vices such as betrayal and treason. Authority/Respect relates to "respect, awe, and admiration toward legitimate authorities" and "virtues related to subordination: respect, duty, and obedience". And Purity/Sanctity relates to the virtues of being "chaste, spiritually minded, pious" and the vices of "lust, gluttony, greed, and anger" (Haidt & Graham, 2007).

[4] "Unity is the motive to care for and support the integrity of in-groups by avoiding or eliminating threats of contamination and providing aid and protection based on need or empathic compassion. Hierarchy is the motive to respect rank in social groups where superiors are entitled to deference and respect but must also lead, guide, direct, and protect subordinates. Equality is the motive for balanced, in-kind reciprocity, equal treatment, equal say, and equal opportunity. Proportionality is the motive for rewards and punishments to be proportionate to merit, benefits to be calibrated to contributions, and judgments to be based on a utilitarian calculus of costs and benefits." (Rai & Fiske, 2011).

[5] "[W]e must abandon the assumption that moral judgments are based on features of actions independent of the social-relational contexts in which they occur (e.g., Did the action cause harm?…Was the action impure?). Rather, we must reconceptualize moral psychology as embedded in our social-relational cognition, such that moral judgments and behaviors emerge out of the specific obligations and transgressions entailed by particular types of social relationships (e.g., Did the action support us against them? Did it go against orders from above? Did you respond in kind?)….[M] oral intuitions are not based on asocial principles of right actions, such as prohibitions against intentionally causing harm…or concerns with 'purity'…Rather, moral intuitions are defined by the particular types of social relationships in which they occur." (Rai & Fiske, 2011).

judgements – for example, to decide with whom to cooperate in future (Krasnow, Delton, Cosmides, & Tooby, 2016), to warn friends and family of uncooperative individuals, to enhance one's reputation as trustworthy or heroic (Barclay, 2016), or to recruit allies to prosecute an offender (Petersen, 2013) – STTM argues instead that the sole function of moral judgements is to provide salient focal points around which people coordinate when taking sides in interpersonal conflicts (DeScioli & Kurzban, 2009, 2013). STTM maintains that a wide range of content, including cooperative rules, can fulfil this function.

Thus MAC makes predictions about the content and structure of morality that are more extensive and detailed than those of previous theories. For the remainder of this paper we will focus on testing MAC's predictions against those of the most well-developed theory – MFT – and return to the implications of our findings for the other theories in the general discussion.

Previous empirical research provides some support for MAC's predictions about the content and structure of morality.

3.1. The content of morality

With regard to content, an analysis of the historical ethnographic records of 60 societies found that the moral valence of these seven cooperative behaviours was uniformly positive, and that there is evidence for the majority of these cooperative moral values in the majority of cultures, in all regions of the world (Curry et al., 2019). Research on more contemporary populations paints a similar picture. First, a survey of family values involving student samples from 30 countries (Byrne & van de Vijver, 2014; Georgas, Berry, Van de Vijver, Kagitçibasi, & Poortinga, 2006) and responses to items in the World Values Survey, conducted in over 65 societies (Inglehart & Baker, 2000), indicate that 'helping kin' is widely considered to be morally good. Second, responses from internet samples to the Ingroup items in the Moral Foundations Questionnaire (Graham et al., 2011), and responses from student samples in 20 countries to items from the Schwartz Basic Values Survey (Schwartz, 1992) both indicate that 'helping your group' is widely considered to be morally good. Third, endorsement of the norms of positive and negative reciprocity in student samples (Eisenberger, Lynch, Aselage, & Rohdieck, 2004), in Britain and Italy (Perugini, Gallucci, Presaghi, & Ercolani, 2003), and responses to some items in the Values in Action Inventory of Strengths in 54 countries (Park, Peterson, & Seligman, 2006; Peterson & Seligman, 2004) and Schwartz's Values Scale (Schwartz, 1992) indicate that 'reciprocity' is widely considered to be morally good. Fourth, investigations into the concept of honour, among students in the US and Turkey (Cross et al., 2014) indicate that various hawkish traits such as bravery are considered to be morally good. Fifth, responses to Authority items in the Moral Foundations Questionnaire (Graham et al., 2011), and to items from the Schwartz Basic Values Survey (Schwartz, 1992) indicate that 'respecting superiors' is widely considered to be morally good. Sixth, responses to items in the Merit Principle Scale in student samples (Davey, Bobocel, Son Hing, & Zanna, 1999) indicate that 'dividing disputed resources' is considered to be morally good. And seventh, responses to items in the World Values Survey (reported in Weeden & Kurzban, 2013) indicate that 'respecting property' is widely considered to be morally good.

However, previous research has not provided a full test of MAC's predictions about the content of morality; no previous study has investigated the moral relevance of all seven forms of cooperative behaviour in a single, contemporary, representative sample. Instead, the studies reviewed above have measured different aspects of morality, in different ways; the scales they employ typically measure something other than the moral relevance (or valence) of cooperation (for example, they ask whether a person or a society possesses a particular trait, rather than whether the trait is moral); and the samples they use are typically composed only of students.

3.2. The structure of morality

With regard to structure, no previous research has investigated MAC's prediction that these seven different types of cooperation will give rise to distinct domains of morality. This is because no previous attempts to map the moral domain – even those that have argued that the function of morality is to promote cooperation – have been guided by the mathematics of cooperation reviewed above, and hence none contain all of the domains predicted by MAC (Curry, 2016).

Nevertheless, despite its limitations, it is possible to ask whether previous work using the Moral Foundations Questionnaire (MFQ) supports MAC's predictions where the two theories overlap. Here the evidence is mixed. Factor analysis has provided only limited support for MFT's five-factor model. The original exploratory factor analysis of data collected using the MFQ suggested a two-factor model (Table 2 in Graham et al., 2011). Confirmatory factor analysis of this data suggested that MFT's five-factor model provided a better fit; but the size of the improvement was marginal, and more importantly, none of the resulting five-factor models exhibited a conventionally 'acceptable' model fit (CFIs ≤ 0.88; Table 10; Graham et al., 2011). Subsequent independent replications in Italy (CFI = 0.88; Bobbio, Nencini, & Sarrica, 2011), New Zealand (CFI = 0.83; Davies, Sibley, & Liu, 2014), Korea (CFI = 0.68; Glover et al., 2014), Sweden (CFI = 0.68; Nilsson & Erlandsson, 2015), and Turkey (CFI = 0.78; Yilmaz, Harma, Bahcekapili, & Cesur, 2016), as well as a 27 country study using the short-form MFQ (CFIs ≤ 0.70; Iurino & Saucier, submitted), all suggest a similar pattern. For this reason, an alternative two-factor model – consisting of an 'individualising' domain of Care and Fairness, and the 'binding' domain of Ingroup, Authority and Purity – is typically used in research (for example, see: Lewis & Bates, 2010; Smith, Alford, Hibbing, Martin, & Hatemi, 2016).

Thus empirical research with the MFQ does not support MAC's prediction that group, deference and fairness will be distinct domains; but it does support MAC's prediction that domains *not* tied to specific forms of cooperation – namely Care and Purity – will *not* constitute distinct domains.

However, it is not clear whether these findings indicate a problem with the cooperative approach to morality in general, or merely a problem with the way that it has been operationalised and measured in Moral Foundations Theory and the MFQ. After all, proponents of MFT have acknowledged that the original list of foundations was somewhat "arbitrary" (p. 107), based on a limited review of only "five books and articles" (p. 107); that this list was never meant to be "exhaustive" (p. 104); and that they "do not know how many moral foundations there really are" (p. 58). And they have positively encouraged research that could "demonstrate the existence of an additional foundation, or show that any of the current five foundations should be merged or eliminated" (Graham et al., 2013, p. 99).

And so, in order to test MAC's predictions – that there will be three additional domains (Family, Heroism, Property), that Reciprocity should not be merged with Fairness; and that Care and Purity should be eliminated – and to overcome the limitations of MFT and the MFQ, we set out to develop a

new measure of morality, the 'Morality-as-Cooperation Questionnaire'.[6]

4. Study 1: Developing a Morality-as-Cooperation Questionnaire (MAC-Q)

In order to test MAC's predictions regarding the content and structure of morality, we used the theories of cooperation reviewed above to develop a self-report measure of moral attitudes to seven types of cooperative behaviour (helping your family, helping your group, reciprocating costs and benefits, being brave, deferring to authority, dividing disputed resources, and respecting property), used it to gather data from a large sample representative of the UK adult population, and submitted the results to factor analysis.

4.1. Methods

In order to measure moral attitudes to the seven types of cooperative behaviour, we followed the MFQ in employing two scales, addressing Relevance and Judgement, each of which is composed of multiple three-item subscales reflecting each of the proposed moral domains (MFQ; Graham et al., 2011).[7]

The Relevance scale was originally developed in order to test whether, in addition to 'liberal' concerns such as care and fairness, 'conservative' concerns such as authority, group loyalty, and purity are also considered relevant to morality (Graham, Haidt, & Nosek, 2009; Haidt & Graham, 2007). Here we use the scale to test whether the types of cooperative behaviour envisaged by MAC (especially the new domains of Family, Reciprocity, Heroism and Property) are considered morally relevant, and also whether they represent distinct domains. The Relevance scale asks participants "When you decide whether something is right or wrong, to what extent are the following considerations relevant to your thinking?". Sample items from our version included "Whether or not someone did what they had agreed to do", and "Whether or not someone kept something that didn't belong to them". Participants responded using an animated slider on a visual analogue scale marked 0–100, and labelled from left-to-right "Not at all Relevant, Not very relevant, Slightly relevant, Somewhat relevant, Very relevant, Extremely relevant".

The Judgement scale was originally developed in order to overcome a perceived limitation of the Relevance scale, namely that it might assess second-order views about how one makes moral judgements, rather than the first-order moral judgements themselves. According to Graham et al. (2009), relevance "does not necessarily measure how people actually make moral judgements", hence

Judgement items were introduced to provide "more contextualized and concrete items that could more strongly trigger the sorts of moral intuitions that are said to play an important role in moral judgement". Here we use the scale to provide an additional test of whether the types of cooperative behaviour envisaged by MAC represent distinct domains. The Judgement scale asks participants "To what extent do you agree with the following statements?". Sample items included "You should always put the interests of your family first", and "Courage in the face of adversity is the most admirable trait". Again, participants responded using an animated slider on a 0–100 visual analogue scale, which was labelled "Strongly Disagree, Disagree, Neither, Agree, Strongly Agree".

We generated items for the new MAC-Q scales based on a comprehensive review of the game-theoretic, evolutionary, psychological, and philosophical literature on cooperation and morality outlined above (Curry, 2005, 2016; Curry et al., 2019). We focussed on the principal moral value in each of the seven hypothesised domains of moral value. For example, for Hawkishness, we focussed on heroism, as opposed to fortitude, generosity or wit. And we aimed to produce items that could, in principle, be used and understood by anyone at any time – the items focussed on interpersonal relationships, and avoided any mention of modern technology or governments. The two sets of items were pretested, to identify heavily skewed or bimodal items. This resulted in 57 Moral Relevance items and 97 Moral Judgement items (available at https://osf.io/w5ad8/).

4.1.1. Participants & procedure

Participants were recruited via a market polling firm (PureProfile). An invitation email was sent to a sample of adults over 18 years-of-age living in the UK and being primary speakers of English. The study was made available on the PureProfile website from 3 to 14 September 2015.

Respondents were provided with information on the study aims and methods. If they provided informed consent, participants were directed to the study site. Participants were told that the goal of the study was to investigate people's sense of right and wrong – described as "what people think of as morally good versus morally bad; moral versus immoral; ethical versus unethical; praiseworthy versus blameworthy". They then completed versions of the Relevance and Judgement scales. The two scales, and items therein, were presented in randomised order. Participants then provided basic demographic data (including age, sex, nationality). Participants could participate only once. Participation took around 12 min, for which participants were paid the equivalent of £0.5.

Responses collected via anonymous internet-based surveys (such as this) depend on participants' sustained effort and engagement, and so are vulnerable to inattention and careless-responding, which affect data quality and the integrity of results (Meade & Craig, 2012). To detect careless responses we used a direct, instructed question as an attention check (Desimone, Harms, & Desimone, 2015).[8] The attention-check item appeared after the main MAC-Q and MFQ item sets, as careless responding increases with survey length and participation time. We also applied further quality controls to remove participants who: did not complete the survey or did not provide demographic data; or who provided responses suggesting spoiled or inauthentic answers (for example, always rating the items 0, 50 or 100, including those that were reverse-coded).

---

[6] OSC conceived of the study, secured funding, wrote the items, collected data, contributed to preliminary analyses, and took the lead in writing the report. MJC wrote the items, collected data, conducted preliminary analyses, and contributed to writing the report. CJV conducted the main analyses, and contributed to writing the report.

[7] For Study 1, we aimed for a sample size of 10 cases per item (10 * 154), though the procedures are reliably robust with > 1000 cases (MacCallum, Widaman, Zhang, & Hong, 1999). Similarly, for Study 2 we aimed for 10 cases per item (10 * 30). For Study 3a we aimed for 10 cases per free parameter (10 * 42). The resulting sample sizes exceed requirements for three-indicator factors (Wolf, Harrington, Clark, & Miller, 2013). For Study 3b the sample size was determined by the number of participants who accepted the invitation complete the follow-up. All materials used to generate the data reported here are either in the public domain or else available on the Open Science Framework page for this paper (osf.io/w5ad8). We have listed all the materials used but not reported here. And all R and MPlus syntaxes required to reproduce the reported results are also available on the OSF page. Regrettably, we are unable to share the data, because the participant consent form included a form of words that inadvertently precluded data sharing. We have now changed the wording of our standard informed consent form to avoid this problem in future.

---

[8] The text of the attention check was as follows: "Hobbies: Everyone has hobbies. Nevertheless, we would like you to skip this question to show that you are reading carefully. Just click 'next'. Do not click any of the buttons corresponding to bike riding, hiking, swimming, playing sports, reading or watching TV."

4.2. Results

2396 people initially accessed the survey. After removing inattentive (544), incomplete (151), and spoiled (305) responses the final sample consisted of 1392 UK working-age adults (628 males, 763 females, 1 transgender; age M = 47.14, SD = 11.14).

4.2.1. Selecting items for analysis

Our initial item pool contained an oversampling of items per subscale. Accordingly, multicollinearity prevented analysis of the entire pool of items, as indicated by the low determinant of the correlation matrices for the Relevance (det = 4.80 * 10^{-17}) and Judgement items (det = 1.45 * 10^{-19}). This suggests substantial redundancy amongst the items. We therefore reduced the total item pool to three items per subscale; the minimum required to estimate latent variables (Kline, 2005). We retained those items which most closely captured the essence of each domain, based on theoretical consistency and empirical considerations – namely, we examined the correlation of each item with the average of all other items within its subscale, excluding the item itself (item-total correlation without item). This resulted in a reduced item pool of 21 items for Relevance and 21 for Judgement. The determinant for Relevance still fell slightly below the threshold of 1 * 10^{-5} (det = 8.15 * 10^{-6}), but no further problems were encountered during analysis. The determinant for Judgement was acceptable (det = 1.77 * 10^{-3}). The full text of the final sets of Relevance and Judgement items are given in Appendices A and B.

4.2.2. Content

Descriptives for the Relevance and Judgement items are given in Tables S1 and S2. Descriptives for the seven Relevance and Judgement subscales, are given in Table 2. Ratings for the Relevance items and subscales ranged from 'somewhat' to 'very' relevant to morality.

4.2.3. Structure

4.2.3.1. Internal consistency.
Cronbach's alphas for the seven subscales (Table 2) ranged from 0.76 to 0.86 for Relevance, and 0.53 to 0.83 for Judgement.

4.2.3.2. Exploratory factor analysis.
Given that the Relevance and Judgement scales were developed to measure two aspects of moral psychology thought to be different, we began by analysing each separately, before attempting to combine them.

Although we had *a priori* hypotheses about the factor structure, we first conducted exploratory factor analyses, on each scale, to see whether the hypothesised structure emerged from the data. The Kaiser-Meyer-Olkin index indicated superb sampling adequacy for Relevance (KMO = 0.95, individual items: 0.91–0.97), and meritorious sampling adequacy for Judgement (KMO = 0.85, individual items: 0.62–0.92). Bartlett's test of sphericity indicated that neither of the correlation matrices were identity matrices, moral Relevance χ^2(210) = 16207.75, p < .001, and Judgement χ^2(210) = 8766.97, p < .001. The data are thus fit for factor analysis.

Parallel analysis (Horn, 1965) revealed that, for both the moral Relevance and Judgement data, seven factors had eigenvalues greater than those derived from randomly generated data.[9] We extracted these seven factors using factor analysis with ML estimation and oblimin rotation to allow factors to correlate. For Relevance, these factors explained 63% of variance in participants' responses (explained variance per factor 7.22–11.15%). For Judgement, these factors explained 50% of variance in participants' responses (explained variance per factor 5.24–9.84%).

The pattern matrices of the resulting factor solutions for the Relevance and Judgement scales are given in Tables S3 and S4. For both Relevance and Judgement, the resulting seven factors clearly corresponded to the seven hypothesised moral domains. All items loaded highest on their corresponding factor, with an average factor loading of 0.69 for Relevance (ranging from 0.45 to 0.84), and an average of 0.63 for Judgement (ranging from 0.42 to 0.82). Cross-loadings were all smaller than 0.24 (absolute value), which is negligible.

4.2.3.3. Confirmatory factor analysis.
We conducted confirmatory factor analysis in order to measure the objective fit of our seven-factor models to the data (Table S5).[10] MAC's seven-factor model was found to have a 'good' (RMSEA), 'acceptable' (CFI), and 'good' (SRMR) fit for the Relevance data, and a 'mediocre' (RMSEA), 'unacceptable' (CFI), and 'good' (SRMR) fit for the Judgement data.

4.2.4. Combining the scales

Whether and to what extent the Relevance and Judgement scales differ, and whether it is necessary to account for potential discrepancies between them to make a unified scale, has never been explicitly tested. Here we remedy this by comparing a model that does not take any difference between Relevance and Judgement into account, with two alternative models that do.

We began by examining the mean score correlations before model estimation (and latent variable correlations derived from the final model). These revealed that most subscales were moderately correlated, but Fairness and Property were not (Table S6), and that it would be important to take this discrepancy into account when combining the scales.

First, we considered the "simple domains" approach taken by MFT/MFQ, whereby Relevance and Judgement items are taken together as indicators of the seven underlying moral domains. Second, we considered a "different but related" approach, whereby separate latent variables are estimated for each of the seven Relevance and Judgement subscales, and these latent variables are allowed to correlate freely. Third, we considered a "multi-trait, multi-method" approach (Widaman, 1985), which separates the variance of each questionnaire item into two components: A "trait" component, and a "method" component. In this case, the "trait" components are seven latent constructs, reflecting the moral domains. Two "method" components capture variance introduced by the two questionnaire versions: Relevance versus Judgement. The model thus reflects the assumption that there are seven underlying domains of morality (traits), which might be tapped in slightly different ways by the Relevance and the Judgement questionnaires. This allows us to estimate people's scores on the moral domains, controlling for variance relating to the two questionnaire versions. Results of these three models are given in Table S7.[11] The

---

[9] According to contemporary psychometric literature (Dinno, 2009), Horn's (1965) parallel analysis is the preferred method for determining the number of factors to retain.

[10] All analyses were conducted in Mplus Version 7 (Muthén & Muthén, 1998–2012). Per the developers' recommendation, we used robust maximum-likelihood estimation, which yields Satorra-Bentler scaled χ^2 values to account for potential non-normality (Satorra, 2000). Model fit was evaluated using three absolute fit indices, namely the Root Mean Square Error of Approximation (RMSEA, values < 0.01, 0.05 and 0.08 are considered to indicate excellent, good, and mediocre fit; MacCallum, Browne, & Sugawara, 1996), Comparative Fit Index (CFI, acceptable fit > 0.9; Bentler & Bonett, 1980), and Standardized Root Mean Square Residual (SRMR, good fit < 0.08; Hu & Bentler, 1999).

[11] In addition to measures of objective fit, we also provide three comparative fit indices. First, the Akaike Information Criterion (AIC), second, sample-size adjusted Bayesian Information Criterion (saBIC), which can be used to compare non-nested models (Kline, 2005). Lower values on these indices indicate better fit. And third, Satorra-Bentler scaled chi-square difference tests, which indicate whether the fit of one model is significantly different from the fit of another (p < .05) (Satorra & Bentler, 2010).

Table 2
MAC-Q subscales: Means, standard deviations & alphas.

	mean	sd	sk	sk/2*se	ku	ku/2*se	α	Interpretation
Relevance								
Family	67.02	18.56	−0.63	−4.79	0.18	0.67	0.86	Good
Group	59.77	18.43	−0.58	−4.41	0.17	0.67	0.86	Good
Reciprocity	66.45	18.01	−0.72	−5.49	0.53	2.02	0.83	Good
Heroism	61.84	19.00	−0.55	−4.22	0.13	0.51	0.84	Good
Deference	53.89	19.14	−0.35	−2.66	−0.22	−0.85	0.80	Good
Fairness	56.47	18.20	−0.40	−3.02	−0.12	−0.45	0.76	Acceptable
Property	65.53	19.04	−0.64	−4.89	0.18	0.70	0.80	Good
Judgment								
Family	67.76	17.50	−0.49	−3.71	0.17	0.63	0.83	Good
Group	64.64	14.82	−0.51	−3.85	0.77	2.92	0.75	Acceptable
Reciprocity	72.12	12.71	−0.35	−2.67	0.58	2.23	0.68	Questionable
Heroism	66.17	17.50	−0.34	−2.60	−0.12	−0.46	0.71	Acceptable
Deference	54.71	17.82	−0.37	−2.82	−0.13	−0.51	0.69	Questionable
Fairness	70.43	16.85	−0.66	−5.02	0.69	2.61	0.66	Questionable
Property	61.22	16.78	−0.05	−0.41	−0.23	−0.87	0.53	Poor

Note. Range for all items is 0–100.

fit of the Simple Domains model was 'mediocre' (RMSEA), 'unacceptable' (CFI), and 'not good' (SRMR). The fit of the Different But Related model was 'good' (RMSEA), 'acceptable' (CFI), and 'good' (SRMR). And the fit of the Multi-Trait Multi-Method model was 'good' (RMSEA), 'acceptable' (CFI), and 'good' (SRMR). Satorra-Bentler scaled chi-square difference tests (Satorra & Bentler, 2010) indicated that the fit of all three models differed significantly from one another. Thus, we can conclude that: the Simple Domains model was not supported; the two models that allowed potential discrepancies between Relevance and Judgement ratings were better supported; and of these two models, the Different but Related achieved the best fit.

4.3. Discussion

The results support MAC's prediction about the content of morality, namely that each of the seven types of cooperation would be considered morally relevant. Indeed, the four new domains uniquely predicted by MAC (family, reciprocity, heroism, property) were considered *more* relevant than domains shared with MFT (group, deference, fairness).

The results also support MAC's predictions regarding the structure of morality, namely that each of the seven types of cooperation would constitute a distinct domain. Exploratory factor analysis delivered the predicted seven factors, on both the Relevance and Judgement scales. And confirmatory factor analysis demonstrated that the seven factor model was a good fit for both scales, separately and combined. The results also suggest caution when combining the scales: simply aggregating the two methodologically distinct Relevance and Judgement scales is inferior to a model that accounts for potential discrepancies between them.

Thus, as uniquely predicted by MAC, the MAC-Q successfully identifies novel moral domains relating to family, reciprocity (as opposed to fairness), heroism, and property, thereby providing evidence of domains not countenanced by previous theories. And the MAC-Q also succeeds in distinguishing moral domains relating to groups, deference and fairness not distinguished by previous measures. In this way, Study 1 provides a first glimpse of a larger, higher resolution map of the moral domain.

5. Study 2: Testing Moral Foundations (MFT/MFQ)

How does the MAC-Q compare to the MFQ? MFT suggests a five-factor model of morality; however, previous research with the MFQ has yielded mixed results. Here we conduct an indepen-

dent test of MFT's predictions. The results also allow us to investigate whether MAC-Q provides a good measure of the domains recognised by both theories (regarding groups, deference, and fairness); and to test MAC's prediction that domains not related to specific forms of cooperation (Care and Purity) would not constitute distinct domains.

5.1. Methods

Methods were identical to Study 1, with the following exceptions. We used the 30-item MFQ, with the original 1–6 scale (Graham et al., 2011). And the data were gathered during the early piloting and pre-testing phase of candidate MAC-Q items, on three separate occasions: from the 3rd to the 4th of December 2013; and the 11th of April to the 11th May, and the 16th to the 18th of September 2014. (Participants also completed a short-form version of the Big Five Inventory, not reported here (Rammstedt & John, 2007).

5.2. Results

1467 accessed the survey online. After removing inattentive (111), incomplete (67) and spoiled (247) responses the final sample consisted of 1042 UK working-age adults (541 males, 499 females, 2 transgender; age, Mean = 48.06 years, SD = 13.94).

5.2.1. Content

Descriptives for the MFQ Relevance and Judgement items given in Tables S8 and S9. Descriptives for the five subscales, for the Relevance and Judgement scales, are given in Table 3. Ratings for the Relevance items ranged from 'slightly' ('conforming to tradition', 'acting in a way that God would approve') to 'very' relevant to morality. Ratings for the Relevance items ranged from 'somewhat' to 'very' relevant (Care).

5.2.2. Structure

5.2.2.1. Internal consistency. Cronbach's alpha for the five subscales (Table 3) ranged from 0.60 to 0.75 for Relevance, and 0.28 to 0.68 for Judgement.

5.2.2.2. Exploratory factor analysis. Again, although we had *a priori* hypotheses about the factor structure, we first conducted exploratory factor analyses, on each scale separately, to see whether the hypothesised structure emerged from the data. The determinants for Relevance and Judgement were both below the threshold of

Table 3
MFQ subscales: Means, standard deviations & alphas.

	mean	sd	sk	sk/2*se	ku	ku/2*se	α	
Relevance								
Care	4.63	0.86	−0.43	−2.83	0.16	0.52	0.75	Acceptable
Fairness	4.46	0.86	−0.29	−1.88	−0.02	−0.08	0.75	Acceptable
Ingroup	3.77	0.99	−0.13	−0.85	−0.26	−0.85	0.74	Acceptable
Authority	3.89	0.90	−0.20	−1.32	0.14	0.47	0.65	Questionable
Purity	3.80	1.03	−0.12	−0.82	−0.15	−0.49	0.60	Questionable
Judgment								
Care	4.70	0.86	−0.47	−3.08	−0.33	−1.10	0.47	Unacceptable
Fairness	4.39	0.73	0.04	0.28	−0.27	−0.90	0.28	Unacceptable
Ingroup	3.92	0.88	−0.13	−0.86	0.10	0.33	0.51	Poor
Authority	4.42	0.90	−0.45	−3.00	0.12	0.39	0.54	Poor
Purity	4.02	1.07	−0.38	−2.51	−0.20	−0.66	0.68	Questionable

Note. Range for all subscales is 1–6.

$1 * 10^{-5}$ (det = $3.26 * 10^{-3}$ and $5.24 * 10^{-2}$, respectively). The Kaiser-Meyer-Olkin index indicated superb sampling adequacy for Relevance (KMO = 0.92, individual items: 0.89–0.95), and meritorious sampling adequacy for Judgement (KMO = 0.84, individual items: 0.58–0.88). Bartlett's test of sphericity indicated that neither of the correlation matrices were identity matrices, Relevance $\chi^2(105) = 5927.70$, $p < .001$, and Judgement $\chi^2(105) = 3053.52$, $p < .001$. The data are thus fit for factor analysis.

For the Relevance scale, parallel analysis revealed that only three factors had eigenvalues greater than those derived from randomly generated data. These factors explained 48% of variance in participants' responses (explained variance per factor ranged from 7% to 21%). The pattern matrix of the resulting factor solution for the Relevance scale is displayed in Table S10. The three factors can be interpreted as: (i) Care/Fairness, (ii) Ingroup/Authority, and (iii) Disgust (consisting of a single item). The average factor loading was 0.60 (ranging from 0.33 to 0.76); cross-loadings were all smaller than 0.31 (absolute values). Extracting five factors did not reveal the hypothesised domains (also Table S10).

For the Judgement scale, parallel analysis revealed that five factors had eigenvalues greater than those derived from randomly generated data. These factors explained 40% of variance in participants' responses (explained variance per factor ranged from 5% to 10%). The pattern matrix of the resulting factor solution for the Judgement scale is given in Table S11. The resulting five factors can be interpreted as: (i) Care/Fairness, (ii) Purity, (iii) Patriotic Authority, (iv) Family and (v) Gender Roles (consisting of a single item). These do not correspond closely to the five hypothesised foundations. The average absolute factor loading was 0.54 (ranging from 0.34 to 0.71); cross-loadings were all smaller than 0.25 (absolute values).

5.2.2.3. Confirmatory factor analysis. We conducted confirmatory factor analysis in order to measure the objective fit of MFT's five-factor model to the data (Table S12). The five-factor model provides a 'mediocre' (RMSEA), 'acceptable' (CFI), and 'good' (SRMR) fit to the Relevance data, and a 'mediocre' (RMSEA), 'unacceptable' (CFI), and 'good' (SRMR) fit to the Judgement data.

5.2.3. Combining the scales
Again, we investigated whether and to what extent these two scales form one unified measure, or whether it is necessary to account for potential discrepancies between them. Mean score correlations before model estimation (and latent variable correlations derived from the final model) revealed that most subscales were moderately correlated (Table S13). Nevertheless, we compared the same three models as before, as well as a fourth Hierarchical model (Graham et al., 2011), which represents the two super-factors of individualising and binding (Table S14).

The fit of the Simple Domains model was 'mediocre' (RMSEA), 'unacceptable' (CFI), and 'not good' (SRMR). The fit of the Different But Related model was 'mediocre' (RMSEA), 'unacceptable' (CFI), and 'good' (SRMR). The fit of the Multi-Trait Multi-Method model was 'mediocre' (RMSEA), 'acceptable' (CFI), and 'good' (SRMR). The fit of the Hierarchical model was 'mediocre' (RMSEA), 'unacceptable' (CFI), and 'not good' (SRMR). Satorra-Bentler scaled chi-square difference tests indicated that the fit of the Multi-Trait Multi-Method model did not differ significantly from that of the Different but Related model, but both of these models fit significantly better than the five-factor Simple Domains model commonly used in the literature, and better than the Hierarchical model. Thus the data did not support the Simple Domains model, or the Hierarchical model. The best fitting models were those that allowed for potential discrepancies between Relevance and Judgement ratings – the Multi-Trait Multi-Method and Different But Related models.

5.3. Discussion

The results of this study support MFT's claim that the contents of the five foundations are considered morally relevant. As such, the results support MAC's prediction that cooperative behaviour relating to groups, deference and fairness (as well as the general category of 'care') will be regarded as morally relevant; but they appear to contradict MAC's prediction that 'purity' will not be considered morally relevant.

However, the results do not support MFT's proposed five-factor structure of morality. Consistent with previous research, the overall internal reliability of the scale was low.[12] Exploratory factor analysis did not yield the predicted five factors, on either Relevance or Judgement scales. (Nor did it reliably reveal the two 'individualising' and 'binding' factors.) Confirmatory analysis demonstrated that none of the five-factor models achieved a good fit on all criteria, either separately or when combined. Specifically, the MFQ was not able to distinguish between moral attitudes relating to groups and deference; Ingroup and Authority items loaded on the same factor. Nor was the MFQ able to distinguish attitudes to fairness; Fairness items loaded on the same factor as Care. Neither Care nor Purity emerged as distinct factors on the Relevance scale, only Purity was a distinct factor on the Judgement scale.

Thus MFT's predictions were not supported. The MFQ was not able to identify the domains uniquely predicted by MFT (Care

---

[12] The internal consistency (Cronbach's alpha) of the MFQ's foundation-specific subscales has been low, for Relevance (range = 0.39–0.76 in Graham et al., 2009; range = 0.65–0.71 in Graham et al., 2011) and especially for Judgement (range = 0.24–0.74 in Graham et al., 2009; range = 0.40–0.75 in Graham et al., 2011).

and Purity), and the MFQ not able to distinguish the domains (groups, deference and fairness) predicted by both MFT and MAC.

Taken together, the results of Studies 1 and 2 provide further support for MAC. First, as MAC predicts, cooperative behaviour is considered relevant to morality (although the relevance of 'purity' remains to be explained), and the seven distinct types of cooperation give rise to seven distinct moral domains. Second, as MAC also predicts, behaviours not tied to a specific form of cooperation ('care' and 'purity') do not reliably emerge as distinct domains. Third, the MAC-Q distinguishes distinct domains relating to groups, deference and fairness, whereas the MFQ does not. And fourth, the MAC-Q's breadth of conceptual coverage, internal consistency, factor structure, and model fit, are all superior to those of the MFQ.

Of course, in order to provide a more robust test of MAC's predictions, of the replicability of MAC's proposed model, and to establish the relative utility of the MAC-Q, it will be necessary to examine the psychometric properties of the final scale, and compare them to the MFQ.

6. Study 3a: Confirming and Validating the MAC-Q

In order to replicate and confirm a further test of MAC's predictions regarding the content and structure of morality, to examine the psychometric properties of the final scale (including its external and predictive validity), and to investigate how these compare to those of the MFQ, we gathered a further round of data using the MAC-Q, the MFQ, and a range of external scales and measures addressing related constructs.

6.1. Methods

Methods were identical to Study 1, with the following exceptions. Data were gathered 27 June to 12 July 2016. Participants were recruited via a market polling firm (Qualtrics.com) in the US. Participation took around 27 min, for which participants were paid the equivalent of $0.75. We used the 42-item MAC-Q, the 30-item MFQ, and a series of domain-related criteria scales to assess divergent and convergent validity.

MAC Kinship criterion scales were the Relations subscale of the Family Values Scale (FVS; Byrne & van de Vijver, 2014), and the Family Security item from the Schwartz's Basic Values Scale (SVS; Schwartz, 1992). MAC Mutualism (and MFQ Ingroup) criterion scales were the Citizen/Teamwork subscale of the Virtues in Action Scale (VIA; Goldberg et al., 2006; Peterson & Seligman, 2004), the Loyalty and National Security items from the SVS, and pictorial measures of Identify Fusion with Community and Country (Gómez et al., 2011). MAC Reciprocity (and MFQ Fairness) criterion scales were the Positive subscale of the Personal Norm of Reciprocity Scale (PNR; Perugini et al., 2003), the Kindness/Generosity subscale of the VIA, and the kindness the Reciprocation item from the SVS. MAC Heroism criterion scales were the Valor/Bravery/Courage subscale of the VIA, and the Power subscale of the VIA. MAC Deference (and MFQ Authority) criterion scales were the Conformity and Tradition subscales of the SVS, and the Modesty/Humility subscale of the VIA. MAC Fairness (and MFQ Fairness) criterion scales were the Equity/Fairness subscale of the VIA, and the Equality and Social Justice items from the SVS. MAC Property criterion scale was the Wealth item from the SVS. In addition, we also collected data on two criteria scales related to the additional MFQ domains. Care criterion scales were items from the Benevolence subscale of the SVS. And the Purity criterion scales were the self-discipline, clean, and devout items on the SVS.

For predictive validity, we also included a new quasi-objective measure of cooperative behaviour. The Cooperative Action Scale asked participants how many times they had performed various cooperative actions in past year (Table S32).[13]

6.2. Results

1157 accessed the survey online. After removing inattentive (537), incomplete (70) and spoiled (81) responses the final sample consisted of 469 US working-age adults (238 males, 230 females, 1 transgender; age, *Mean* = 46.82 years, *SD* = 16.78).

6.2.1. MAC-Q content

Descriptives for all MAC-Q Relevance and Judgement items are given in Tables S15 and S16. Descriptives for all seven MAC-Q subscales are given in Table 4. Ratings for the Relevance items and subscales ranged from 'somewhat' to 'very' relevant to morality.

6.2.2. MAC-Q structure

6.2.2.1. Internal consistency. Cronbach's alphas for all seven subscales (Table 4) ranged from 0.76 to 0.88 for Relevance, and 0.64 to 0.86 for Judgement.

6.2.2.2. Exploratory factor analysis. Once again, we conducted exploratory factor analyses, on each scale separately, to see whether the hypothesised structure emerged from the data. The determinant for Relevance fell slightly below the threshold of 1e−5, 3.10e−06. The determinant for Judgement was good, 1.95e−4. The Kaiser-Meyer-Olkin index indicated superb sampling adequacy for Relevance (KMO = 0.93, individual items: 0.87–0.95), and meritorious sampling adequacy for Judgement (KMO = 0.88, individual items: 0.68–0.95). Bartlett's test of sphericity indicated that neither of the correlation matrices were identity matrices, moral Relevance $\chi^2(210)$ = 5836.55, $p < .001$, and Judgement $\chi^2(210)$ = 3920.71, $p < .001$. The data are thus fit for factor analysis.

For the Relevance scale, parallel analysis revealed that *seven* factors had eigenvalues greater than those derived from randomly generated data. These factors explained 66% of variance in participants' responses (explained variance per factor ranged from 7% to 11%). The pattern matrix of the resulting factor solution for the Relevance scale is given in Table S17. The resulting seven factors clearly corresponded to the seven hypothesised moral domains. All items loaded highest on their corresponding factor, with an average factor loading of 0.69 (ranging from 0.45 to 0.84). Cross-loadings were all smaller than 0.24 (absolute value), which is negligible.

For the Judgement scale, parallel analysis revealed that *five* factors had eigenvalues greater than those derived from randomly generated data. These factors explained 51% of variance in participants' responses (explained variance per factor 8–13%). The pattern matrix of the resulting factor solution for the Judgement scale is displayed in Table S18. The five factors can be interpreted as: (i) Family, (ii) Group/Reciprocity, (iii) Martial Virtues, (iv) Fairness and (v) Property. Broadly speaking, mutualism and reciprocity, and heroism and deference, were combined. The average absolute factor loading was 0.58 (ranging from 0.29 to 1.01); cross-loadings were all smaller than 0.40 (absolute values). However, extracting seven factors did reveal the hypothesised domains (also Table S18), with all items loading highest on their corresponding factor, with an average factor loading of 0.60 for Relevance (ranging from 0.34 to 0.87), and cross-loadings smaller than 0.38 (absolute value).

---

[13] We also collected data on: Big 5 (Rammstedt & John, 2007); Social and Economic Conservatism (Everett, 2013); Social Desirability Responding (Reynolds, 1982); hypothetical kidney donation; age, sex, nationality/state, language, religion, religiosity, ethnicity, political affiliation, and voting intention, which we plan to report in subsequent papers.

Table 4
MAC-Q subscales: Means, standard deviations, alphas & retest reliabilities.

	mean	sd	sk	sk/2*se	ku	ku/2*se	α	Interpretation	Retest
Relevance									
Family	73.50	20.27	−1.00	−4.45	1.08	2.39	0.87	Good	0.88
Group	66.90	20.94	−0.82	−3.65	0.61	1.36	0.88	Good	0.89
Reciprocity	76.18	18.04	−0.91	−4.02	0.82	1.82	0.84	Good	0.86
Heroism	67.94	20.53	−0.81	−3.58	0.56	1.24	0.82	Good	0.84
Deference	65.20	20.77	−0.68	−3.01	0.35	0.77	0.79	Acceptable	0.81
Fairness	61.59	20.56	−0.47	−2.10	0.04	0.09	0.76	Acceptable	0.79
Property	70.40	21.96	−0.94	−4.18	0.62	1.37	0.84	Good	0.86
Judgment									
Family	71.01	20.80	−0.77	−3.43	0.31	0.68	0.86	Good	0.87
Group	71.58	16.19	−0.59	−2.60	0.86	1.90	0.73	Acceptable	0.74
Reciprocity	77.24	14.19	−0.44	−1.97	−0.09	−0.20	0.70	Acceptable	0.71
Heroism	73.58	17.72	−0.84	−3.71	0.75	1.66	0.76	Acceptable	0.78
Deference	64.15	19.07	−0.63	−2.79	0.20	0.45	0.71	Acceptable	0.71
Fairness	77.66	16.53	−1.15	−5.11	2.25	5.01	0.64	Questionable	0.66
Property	62.16	22.79	−0.39	−1.73	−0.68	−1.51	0.70	Acceptable	0.71

Note. Range for all items is 0–100.

6.2.2.3. Confirmatory factor analysis.

MAC's seven-factor model was found to have a 'good' (RMSEA), 'acceptable' (CFI), and 'good' (SRMR) fit for the Relevance data, and a 'mediocre' (RMSEA), 'unacceptable' (CFI), and 'good' (SRMR) fit for the Judgement data (Table S19).

6.2.3. Combining the scales

Again, we investigated whether and to what extent these two scales form one unified measure, or whether it is necessary to account for potential discrepancies between them.

Mean score correlations before model estimation (and latent variable correlations derived from the final model) revealed that most subscales were moderately correlated, but Fairness and Property were not (Table S20). This again suggested that it would be necessary to take this discrepancy into account when creating a scale that combined the two Relevance and Judgement scales (Table S21).

The fit of the Simple Domains model was 'unacceptable' (RMSEA), 'unacceptable' (CFI), and 'not good' (SRMR). The fit of the Different But Related model was 'good' (RMSEA), 'acceptable' (CFI), and 'good' (SRMR). The fit of the Multi-Trait Multi-Method model was 'good/mediocre' (RMSEA), 'unacceptable' (CFI), and 'good' (SRMR). Satorra-Bentler scaled chi-square difference tests indicated that the fit of all three models differed significantly from one another. Thus the Simple Domains model was not supported. The two models that allowed potential discrepancies between Relevance and Judgement ratings achieved a significantly better fit on all indices. And of these two models, the Different but Related model achieved the best fit.

6.2.4. MFQ content

Descriptives for all MFQ Relevance and Judgement items are given in Tables S22 and S23. Descriptives for all five MFQ subscales are given in Table 5. Ratings for the Relevance items and subscales ranged from 'somewhat' to 'very' relevant to morality.

6.2.5. MFQ structure

6.2.5.1. Internal consistency.
Cronbach's alphas for the five subscales ranged from 0.66 to 0.77 for Relevance, and 0.44 to 0.75 for Judgement.

6.2.5.2. Exploratory factor analysis.
Again, we first conducted exploratory factor analyses, on each scale separately, to see whether the hypothesised structure emerged from the data. The determinants for Relevance and Judgement were both above the threshold of $1 * 10^{-5}$ (det = $7.63 * 10^{-4}$ and $2.00 * 10^{-2}$, respectively), indicating the correlation matrices were fit for factor analysis.

The Kaiser-Meyer-Olkin index indicated superb sampling adequacy for Relevance (KMO = 0.93, individual items: 0.86–0.95), and meritorious sampling adequacy for Judgement (KMO = 0.85, individual items: 0.70–0.89). Bartlett's test of sphericity indicated that neither of the correlation matrices were identity matrices, Relevance $\chi^2(105)$ = 3317.67, $p < .001$, and Judgement $\chi^2(105)$ = 1810.67, $p < .001$. The data are thus fit for factor analysis.

For the Relevance scale, parallel analysis revealed that only *two* factors had eigenvalues greater than those derived from randomly generated data. These factors explained 51% of variance in participants' responses (explained variance per factor 24% - 27%). The

Table 5
MFQ subscales: Means, standard deviations, alphas & retest reliabilities.

	mean	sd	sk	sk/2*se	ku	ku/2*se	α	Interpetation	Retest
Relevance									
Care	4.77	0.95	−0.99	−4.39	1.38	3.07	0.76	Acceptable	0.78
Fairness	4.73	0.98	−0.89	−3.97	1.16	2.57	0.77	Acceptable	0.80
Ingroup	4.31	1.05	−0.61	−2.69	0.13	0.28	0.75	Acceptable	0.77
Authority	4.31	1.00	−0.60	−2.66	0.42	0.93	0.66	Questionable	0.70
Purity	4.19	1.22	−0.57	−2.51	−0.40	−0.89	0.72	Acceptable	0.74
Judgment									
Care	4.68	0.95	−0.62	−2.75	0.23	0.51	0.49	Unacceptable	0.51
Fairness	4.39	0.87	−0.44	−1.96	0.31	0.69	0.44	Unacceptable	0.46
Ingroup	4.02	1.08	−0.27	−1.19	−0.51	−1.14	0.60	Questionable	0.62
Authority	4.50	0.95	−0.68	−3.01	0.50	1.11	0.56	Poor	0.59
Purity	4.14	1.23	−0.62	−2.73	−0.22	−0.48	0.75	Acceptable	0.75

Note. Range for all subscales is 1–6.

pattern matrix of the resulting factor solution for the Relevance scale is displayed in Table S24. The two factors can be interpreted as: (i) Care/Fairness, (ii) Ingroup/Authority/Purity (with one errant item, Chaos). The average factor loading was 0.65 (ranging from 0.44 to 0.78); cross-loadings were all smaller than 0.39 (absolute values). Extracting five factors did not reveal the hypothesised domains (also Table S24).

For the Judgement scale, parallel analysis revealed that *four* factors had eigenvalues greater than those derived from randomly generated data. These factors explained 43% of variance in participants' responses (explained variance per factor 7–15%). The pattern matrix of the resulting factor solution for the Judgement scale is displayed in Table S30. The two factors can be interpreted as: (i) Care/Government Fairness, (ii) Family/Team, (iii) Martial Virtue/Inheritance and (iv) Purity/Sex Roles. The average absolute factor loading was 0.54 (ranging from 0.35 to 0.78); cross-loadings were all smaller than 0.48 (absolute values). Extracting five factors did not reveal the hypothesised domains (also Table S25).

6.2.5.3. Confirmatory factor analysis. Despite the fact that exploratory factor analysis did not reveal MFT's five-factor structure, we proceeded to conduct confirmatory factor analysis in order to measure the objective fit of MFT's five-factor model to the data (Table S26). The five-factor model provides a 'mediocre' (RMSEA), 'unacceptable' (CFI), and 'good' (SRMR) fit to the Relevance data, and a 'unacceptable' (RMSEA), 'unacceptable' (CFI), and 'good' (SRMR) fit to the Judgement data.

6.2.6. Combining the scales

Again, we investigated whether and to what extent these two scales form one unified measure, or whether it is necessary to account for potential discrepancies between them.

Mean score correlations before model estimation (and latent variable correlations derived from the final model) revealed that most subscales were moderately correlated (Table S27). Nevertheless, we compared the same models as before (Table S28).

The fit of the Simple Domains model was 'unacceptable' (RMSEA), 'unacceptable' (CFI), and 'not good' (SRMR). The fit of the Different But Related model was 'mediocre' (RMSEA), 'unacceptable' (CFI), and 'good' (SRMR). The fit of the Multi-Trait Multi-Method model was 'mediocre' (RMSEA), 'acceptable' (CFI), and 'good' (SRMR). And the fit of the Hierarchical model was 'unacceptable' (RMSEA), 'unacceptable' (CFI), and 'not good' (SRMR). According to Satorra-Bentler scaled chi-square difference tests, the Multi-Trait Multi-Method model fit significantly better than all alternative models. Thus the data did not support the five-factor Simple Domains model, the Different But Related model, or the Hierarchical model. Again, the only model to achieve adequate fit on all indices was one that allowed for potential discrepancies between Relevance and Judgement ratings: the Multi-Trait Multi-Method.

6.2.7. MAC and MFQ external and predictive validity

6.2.7.1. External validity

Correlations between the MAC-Q subscales and the external criterion scales, as well as the average correlation for each criterion group, are shown in Table 6.[14] For MAC-Q Relevance, each subscale was correlated with its own conceptually related group of external scales, with the exception of Property; however, only Family and Deference were the strongest predictors for their own conceptually

---

[14] Correlations were transformed to Fisher's z-values, the z-values were averaged, and the average z-values back-transformed to correlations (Fisher, 1915).

related group of external scales. For MAC-Q Judgement, each subscale was the strongest predictors for its own conceptually related group of external scales, with the exception of Group. Correlations between the MFQ subscales and the external criterion scales are shown in Table 7. For MFQ Relevance, each subscale was the strongest predictors for its own conceptually related group of external scales. For MFQ Judgement, each subscale was correlated with its own conceptually related group of external scales; however, only Fairness, Ingroup and Purity were the strongest predictors for their own conceptually related group of external scales.

Correlations between the Cooperative Actions and MAC-Q and MFQ subscales are shown in Tables S29 and S30.

6.2.7.2. Incremental predictive validity

Following Graham (2011), we used two-step regressions to test whether the seven MAC-Q subscales added incremental predictive validity beyond the five MFQ subscales for the external criteria described above, as well as for the cooperative actions.

Overall, the MAC-Q was a better predictor of Family, Group, Heroism and Property external scales, whereas the MFQ was a better predictor of Deference and Fairness (Table 8). And the MAC-Q was a better predictor than the MFQ of Cooperative Actions in ten cases, equal on three, and worse on one (Table S31).

7. Study 3b: MAC-Q and MFQ Test-retest reliability

To establish and compare the test-retest reliability of the MAC-Q and the MFQ, participants from Study 3a were invited to complete both measures a second time after a one month interval (16–18 August 2016), for which they were paid $1.50. 151 participants completed the survey; after removing those who had been excluded from Study 3a, the final sample consisted of 137 participants (68 males, 69 females; age, *Mean* = 53.10 years, *SD* = 15.84) completed it.

Test-retest Pearson correlations (corrected for attenuation) are provided in Table 4 for the MAC-Q and Table 5 for the MFQ. For the MAC-Q, stability coefficients ranged from 0.79 to 0.89 ('good') for Relevance and 0.66–0.87 ('acceptable' to 'good') for Judgement. For the MFQ, they were 0.70–0.80 ('acceptable') for Relevance, and 0.46–0.75 ('poor' to 'acceptable') for Judgement.

7.1. Discussion

Consistent with the results of Study 1, the results of Study 3 provide further support for MAC's prediction that each of the seven types of cooperation would be considered morally relevant. Again, the four novel domains proposed by MAC (family, reciprocity, heroism, and property) were all considered more relevant than the three domains also proposed by MFT (group, deference, and fairness). The results also provide further support for MAC's predictions that each of the seven types of cooperation would constitute a distinct domain. Exploratory factor analysis delivered the predicted seven factors, on both the Relevance and (to a lesser extent) Judgement scales. And confirmatory factor analysis demonstrated that the seven factor model was a reasonable fit for both scales separately, and a good fit when combined. (And again, simply aggregating the two methodologically distinct Relevance and Judgement scales was shown to be inferior to a model that accounts for potential discrepancies between them.) Thus, as uniquely predicted by MAC, the MAC-Q once again successfully identifies novel moral domains relating to family, reciprocity (as opposed to fairness), heroism, and property; and it succeeds in distinguishing moral domains relating to groups, deference and fairness.

Turning to the MFQ, consistent with Study 2, the results from Study 3 support MFT's claim that the contents of the five foundations

Table 6
Pearson correlations of MAC-Q subscales with external scales and scale items.

External scale criteria group	Relevance Subscales							Judgment Subscales						
	Family	Group	Reciprocity	Heroism	Deference	Fairness	Property	Family	Group	Reciprocity	Heroism	Deference	Fairness	Property
Family														
FVS (Relations)	**0.48**	0.25	0.29	0.35	0.37	0.15	0.20	**0.55**	0.39	0.50	0.46	0.45	0.16	0.23
SVS: Family Security Item	**0.43**	0.31	0.22	0.24	0.25	0.18	0.26	**0.38**	0.34	0.30	0.29	0.23	0.16	0.16
Average	**0.46**	0.28	0.26	0.30	0.31	0.17	0.23	**0.47**	0.37	0.40	0.38	0.34	0.16	0.20
Group														
SVS: Loyal Item	**0.32**	0.29	0.29	0.25	0.23	0.22	0.22	0.27	**0.34**	**0.34**	0.30	0.22	0.19	0.14
SVS: National Security Item	0.31	0.21	0.23	0.27	**0.34**	0.16	0.22	0.32	0.27	0.32	**0.44**	0.33	0.07	0.17
VIA: Citizen/Teamwork	0.08	0.16	0.06	0.11	0.09	**0.17**	0.04	0.10	0.16	**0.19**	0.10	0.18	0.14	−0.30
Community Fusion	0.26	**0.31**	0.16	0.23	0.17	0.13	0.02	0.28	**0.36**	0.23	0.24	0.28	0.13	−0.06
Country Fusion	**0.35**	0.25	0.24	0.32	0.29	0.13	0.07	**0.41**	0.26	0.31	0.40	0.37	0.08	0.06
Average	**0.27**	0.24	0.20	0.24	0.23	0.16	0.11	0.28	0.28	0.28	**0.30**	0.28	0.12	0.00
Reciprocity														
PNR (Positive)	0.32	0.24	0.28	**0.34**	0.21	0.14	0.18	0.27	0.32	**0.56**	0.34	0.17	0.19	0.17
SVS: Reciprocation Item	0.22	**0.25**	0.23	0.22	0.21	**0.25**	0.12	0.15	0.24	**0.32**	0.25	0.22	0.13	−0.01
VIA: Kindess/Generosity	**0.34**	**0.34**	0.23	0.33	**0.34**	0.26	0.16	0.36	0.37	0.35	0.35	**0.42**	0.19	−0.16
Average	0.29	0.28	0.25	**0.30**	0.25	0.22	0.15	0.26	0.31	**0.42**	0.31	0.27	0.17	0.00
Heroism														
VIA: Valor/Bravery/Courage	0.19	**0.26**	0.11	0.21	0.15	0.24	0.05	0.20	0.27	0.21	0.19	0.23	0.17	−**0.32**
Deference														
VIA: Modesty/Humility	**0.29**	0.27	0.20	0.26	**0.29**	0.27	0.14	0.28	0.28	0.30	0.29	**0.38**	0.20	−0.14
SVS: Tradition	0.39	0.38	0.23	0.33	**0.42**	0.28	0.16	0.47	0.42	0.35	0.42	**0.55**	0.14	−0.04
SVS: Conformity	0.43	0.32	0.26	0.32	**0.45**	0.26	0.23	0.44	0.38	0.38	0.41	**0.47**	0.15	0.12
SVS: Power	0.22	0.26	0.09	0.27	0.23	**0.28**	0.01	0.35	0.30	0.22	0.31	**0.44**	0.08	−0.36
Average	0.34	0.31	0.20	0.30	**0.35**	0.27	0.14	0.39	0.35	0.31	0.36	**0.46**	0.14	−0.11
Fairness														
VIA: Equity/Fairness	0.31	**0.35**	0.23	0.25	0.21	0.29	0.20	0.24	0.38	0.37	0.31	0.28	**0.42**	−0.05
SVS: Equality Item	0.22	**0.24**	0.10	0.10	0.13	0.18	0.11	0.16	0.34	0.28	0.20	0.17	**0.55**	0.08
SVS: Social Justice Item	0.30	**0.36**	0.16	0.19	0.19	0.25	0.22	0.20	**0.41**	0.31	0.20	0.18	0.40	0.02
Average	0.28	**0.32**	0.16	0.18	0.18	0.24	0.18	0.20	0.38	0.32	0.24	0.21	**0.46**	0.02
Property														
SVS: Wealth item	0.09	**0.16**	0.02	0.14	0.09	**0.16**	−0.03	0.23	0.17	0.16	0.18	0.24	0.05	−**0.35**
Care														
SVS: Friendship Item	0.23	**0.24**	**0.24**	0.22	0.12	0.15	0.13	0.11	**0.26**	0.25	0.21	0.10	0.19	0.12
Purity														
SVS: Self-discipline Item	**0.26**	0.17	0.18	0.19	0.25	0.13	0.14	0.23	0.21	0.24	**0.27**	0.22	0.08	0.12
SVS: Clean Item	**0.35**	0.27	0.25	0.25	0.30	0.24	0.15	**0.38**	0.32	0.32	0.31	0.36	0.21	0.06
SVS: Devout Item	0.31	0.28	0.16	0.25	**0.35**	0.12	0.10	0.41	0.32	0.24	0.32	**0.46**	0.02	0.03
Average	**0.31**	0.24	0.20	0.23	0.30	0.16	0.13	0.34	0.28	0.27	0.30	**0.35**	0.10	0.07

Note. The largest correlation for each set of scales is shown in bold. MFQ = Moral Foundations Questionnaire; SVS = Schwartz Values Scale; FVS = Family Values Scale; PNR = Personal Norm of Reciprocity; VIA = Virtues in Action.

Table 7
Pearson correlations of MFQ subscales with external scales and scale items.

External scale criteria group	Relevance Subscales					Judgment Subscales				
	Care	Fairness	Ingroup	Authority	Purity	Care	Fairness	Ingroup	Authority	Purity
Family										
FVS (Relations)	0.34	0.28	0.38	0.41	**0.43**	0.31	0.23	0.42	**0.48**	0.40
SVS: Family Security Item	**0.38**	0.30	0.31	0.28	0.31	**0.30**	0.26	0.22	0.27	0.26
Average	0.36	0.29	0.35	0.35	**0.37**	0.31	0.25	0.32	**0.38**	0.33
Group										
SVS: Loyal Item	0.34	0.34	**0.39**	0.34	0.32	**0.32**	0.32	0.23	0.25	0.25
SVS: National Security Item	0.29	0.24	0.39	**0.43**	0.39	0.29	0.26	0.33	**0.37**	0.34
VIA: Citizen/Teamwork	0.12	0.14	**0.19**	0.15	0.12	0.15	**0.27**	0.18	0.15	0.14
Community Fusion	0.10	0.13	**0.26**	0.23	0.25	0.21	0.20	**0.26**	0.09	0.25
Country Fusion	0.15	0.13	**0.34**	0.31	0.34	0.17	0.14	**0.39**	0.29	0.26
Average	0.20	0.20	**0.32**	0.29	0.29	0.23	0.24	**0.28**	0.23	0.25
Reciprocity										
PNR (Positive)	**0.33**	0.24	0.26	0.25	0.26	**0.33**	0.24	0.24	0.29	0.28
SVS: Reciprocation Item	0.22	0.23	**0.32**	0.28	0.27	0.21	**0.25**	0.21	0.23	0.22
VIA: Kindess/Generosity	0.30	0.22	**0.41**	0.40	0.39	0.38	0.38	**0.44**	0.28	0.40
Average	0.28	0.23	**0.33**	0.31	0.31	**0.31**	0.29	0.30	0.27	0.30
Heroism										
VIA: Valor/Bravery/Courage	0.20	0.17	0.23	**0.24**	**0.24**	0.26	**0.34**	0.33	0.20	0.27
Deference										
VIA: Modesty/Humility	0.27	0.26	0.34	**0.40**	0.37	0.34	0.35	0.37	0.31	**0.40**
SVS: Tradition	0.28	0.20	0.46	0.48	**0.55**	0.36	0.36	0.49	0.47	**0.51**
SVS: Conformity	0.35	0.27	0.46	**0.50**	0.46	0.35	0.33	0.42	0.42	**0.43**
SVS: Power	0.06	0.07	**0.33**	0.31	0.30	0.23	0.32	**0.48**	0.28	0.29
Average	0.24	0.20	0.40	**0.43**	0.42	0.32	0.34	**0.44**	0.37	0.41
Fairness										
VIA: Equity/Fairness	0.43	**0.46**	0.35	0.30	0.29	0.46	**0.49**	0.29	0.22	0.26
SVS: Equality Item	0.37	**0.44**	0.20	0.17	0.12	0.38	**0.42**	0.07	0.03	0.07
SVS: Social Justice Item	0.44	**0.46**	0.27	0.25	0.22	0.39	**0.46**	0.11	0.12	0.16
Average	0.41	**0.45**	0.27	0.24	0.21	0.41	**0.46**	0.16	0.12	0.16
Property										
SVS: Wealth item	0.01	0.06	**0.18**	0.15	0.16	0.13	0.20	**0.31**	0.17	0.14
Care										
SVS: Friendship Item	**0.30**	0.27	0.26	0.22	0.19	0.22	**0.23**	0.11	0.10	0.13
Purity										
SVS: Self-discipline Item	0.26	0.18	0.28	**0.35**	0.30	0.25	0.21	0.24	**0.28**	**0.28**
SVS: Clean Item	0.29	0.26	**0.39**	0.35	0.35	0.30	0.29	**0.36**	0.30	0.33
SVS: Devout Item	0.19	0.08	0.36	0.39	**0.59**	0.25	0.19	0.39	0.39	**0.52**
Average	0.25	0.17	0.34	0.36	**0.42**	0.27	0.23	0.33	0.32	**0.38**

Note. The largest correlation for each set of scales is shown in bold. MFQ = Moral Foundations Questionnaire; SVS = Schwartz Values Scale; FVS = Family Values Scale; PNR = Personal Norm of Reciprocity; VIA = Virtues in Action.

are considered morally relevant – thus supporting MAC's prediction that cooperative behaviour relating to groups, deference and fairness (as well as the general category of 'care') will be regarded as morally relevant, but contradicting the prediction that 'purity' will not. However, once again, the results do not support MFT's proposed five-factor structure of morality. Exploratory factor analysis did not deliver the predicted five factors, on either Relevance or Judgement scales. And confirmatory analysis demonstrated that none of the five-factor models achieved a good fit on all criteria, either separately or combined. (Again, simply aggregating the Relevance and Judgement scales was shown to be inferior to a model that accounts for potential discrepancies between them.) The MFQ was not able to identify distinct domains of morality relating to groups, authority and fairness. And neither Care nor Purity emerged as distinct factors. Thus MFT's predictions were not supported. The MFQ was, once again, not able to identify the domains uniquely predicted by MFT (Care and Purity); nor was it able to distinguish the domains (groups, deference and fairness) predicted by both MFT and MAC.

With regard to external and predictive validity, incremental predictive validity, and test-retest reliability, the performance of the MAC-Q and the MFQ was broadly comparable. Although should it be noted that (a) the original MFQ items were prescreened for

their correlations with SVS items, and (b) the results here do not replicate the very high external validity previously reported for the MFQ (Table 7; Graham et al., 2011).

In conclusion, when combined with the results of Studies 1 and 2, the results of Studies 3a and 3b provide further support for MAC's prediction that cooperative behaviour will be considered relevant to morality, and that the seven distinct types of cooperation give rise to seven distinct moral domains (including the novel moral domains of family, reciprocity, heroism and property). Moreover, as MAC also predicts, behaviours not tied to a specific form of cooperation ('care' and 'purity') do not emerge as distinct domains. The MAC-Q yielded and distinguished between domains (groups, deference and fairness) that MFQ did not. And the MAC-Q's breadth of conceptual coverage, internal consistency, factor structure, and model fit, were superior to those of the MFQ.

8. General discussion

Morality-as-Cooperation (MAC) is the theory that morality consists of a collection of solutions to recurrent problems of cooperation. Here we have shown how the mathematics of cooperation – derived from evolutionary biology and nonzero sum game theory – can be used to develop this theory; and by

Table 8

Incremental predictive validity comparisons between MAC-Q and MFQ (Relevance/Judgment) for external scales and items.

External Scales and Items	R^2 MAC-Q-R	R^2 MFQ-R	R^2 adding MAC-Q to MFQ	R^2 adding MFQ to MAC-Q	R^2 MAC-Q-J	R^2 MFQ-J	R^2 adding MAC-Q to MFQ	R^2 adding MFQ to MAC-Q
Family Scales								
FVS (Relations)	**0.29**	0.22	0.12	0.06	**0.44**	0.29	0.19	0.04
SVS: Family Security Item	**0.21**	0.17	0.09	0.05	**0.21**	0.14	0.11	0.04
Group Scales								
SVS: Loyal Item	0.13	**0.18**	0.02	0.08	**0.17**	0.16	0.07	0.05
SVS: National Security Item	0.15	**0.21**	0.02	0.08	**0.24**	0.20	0.08	0.05
VIA: Citizen/Teamwork	**0.05**	0.04	0.04	0.03	**0.16**	0.08	0.10	0.03
Community Fusion	**0.12**	0.08	0.08	0.04	**0.16**	0.11	0.07	0.03
Country Fusion	**0.17**	0.15	0.06	0.04	**0.23**	0.16	0.09	0.01
Reciprocity Scales								
PNR (Positive)	**0.14**	0.12	0.06	0.04	**0.34**	0.16	0.24	0.06
SVS: Reciprocation Item	0.10	**0.11**	0.03	0.05	**0.13**	0.10	0.05	0.03
VIA: Kindess/Generosity	0.17	**0.21**	0.04	0.07	0.26	**0.30**	0.04	0.08
Heroism Scales								
VIA: Valor/Bravery/Courage	**0.10**	0.07	0.07	0.04	**0.19**	0.18	0.09	0.07
Deference Scales								
VIA: Modesty/Humility	0.13	**0.18**	0.03	0.08	0.20	**0.25**	0.03	0.09
SVS: Tradition	0.24	**0.33**	0.06	0.15	0.36	**0.38**	0.06	0.08
SVS: Conformity	0.26	**0.29**	0.05	0.08	**0.31**	0.30	0.07	0.06
SVS: Power	**0.17**	0.16	0.10	0.08	**0.35**	0.26	0.13	0.05
Fairness Scales								
VIA: Equity/Fairness	0.15	**0.24**	0.04	0.13	0.26	**0.31**	0.05	0.10
SVS: Equality Item	0.09	**0.20**	0.05	0.16	**0.32**	0.25	0.12	0.05
SVS: Social Justice Item	0.16	**0.23**	0.06	0.13	0.24	**0.26**	0.07	0.09
Property Scales								
SVS: Wealth item	**0.07**	0.05	0.06	0.04	**0.20**	0.11	0.12	0.02
Care								
SVS: Friendship Item	0.08	**0.11**	0.03	0.05	**0.10**	0.07	0.06	0.02
Purity								
SVS: Self-discipline Item	0.09	**0.13**	0.01	0.06	0.10	**0.13**	0.03	0.05
SVS: Clean Item	0.15	**0.17**	0.04	0.06	**0.20**	0.20	0.04	0.04
SVS: Devout Item	0.17	**0.37**	0.04	0.24	0.27	**0.30**	0.06	0.09

Note. Bold indicates which scale, MAC-Q or MFQ, is the stronger predictor alone.

identifying seven candidate types of cooperative behaviour, we have extended the theory to incorporate and explain more aspects of morality than previous cooperative accounts.

We have also tested MAC's predictions regarding the content and structure of morality, over the course of four studies. Regarding content, the results support the prediction that all seven types of cooperative behaviours – helping kin, helping one's group, reciprocating costs and benefits, displaying 'hawkish' and dove-ish traits, dividing disputed resources, and respecting prior possession – will be considered relevant to morality. And regarding structure, the results support the prediction that there will be distinct moral domains dedicated to family, groups, reciprocity, heroism, deference, fairness and property. In this way, MAC goes beyond previous theories of morality, including MFT, to identify for the first time novel moral domains of morality relating to family, reciprocity, heroism and property.

More specifically, the results support MAC's claim that (contrary to RMT) 'family' can be distinguished from 'group', and (contrary to MFT and RMT) 'reciprocity' can be distinguished from 'fairness'. And the results support MAC's prediction that behaviour *not* tied to specific forms of cooperation ('care' and 'purity') will *not* form distinct moral domains (consistent with RMT, and RMT and TDM, respectively).

These studies have also produced a new scale for the measurement of morality – the MAC-Q – that exhibits broader and more detailed coverage than, and superior psychometric properties to, the previous leading scale. The results also question the routine

combination of Relevance and Judgement scales. As originally anticipated (Graham et al., 2009), the Relevance and Judgement scales seem to measure somewhat disparate aspects of morality. Across three studies, we found consistent evidence indicating that Relevance and Judgement items should not be combined into a common scale without accounting for their differences. Until the reasons for this discrepancy between moral relevance and judgement is understood, we recommend either combining the measures using a MTMM model, as described above, or using one or both scales separately.

8.1. Limitations and future directions

First, the present study tested the general theory of MAC with respect to seven specific types of cooperation. Future research search should test the theory more widely still, using additional examples of cooperative behaviour. These might include 'subcomponents' of the types of cooperation discussed here; for example, 'social exchange' involves not just reciprocity, but also trust, gratitude, guilt, apology and forgiveness. Or it might include novel types of cooperation yet to be discovered or adumbrated by game theory and the behavioural sciences. Such research could extend MAC to other, as yet poorly understood, aspects of morality.

Second, the present study found that even though 'care' and 'purity' did not reliably emerge as unitary domains, the items they contained were nevertheless rated as *relevant* to morality. Future research should aim to explain why. Perhaps, as MAC suggests,

these constructs reflect the operation of proximate mechanisms, such as sympathy and disgust, that contribute to the solution of multiple distinct problems of cooperation.[15]

Third, the present study has successfully 'isolated' seven different types of morality. Future research should investigate how they interact. For example, MAC predicts that having to choose between alternative, incompatible cooperative courses of action will give rise to moral dilemmas. Should you tend to your ailing mother, or go off to fight for your country (Sartre, 1946/1973)? MAC also predicts that when one cooperative opportunity is pursued at the expense of some larger more valuable opportunity ('the greater good'), the former will be regarded as (relatively) morally bad (Muthukrishna, Francois, Pourahmadi, & Henrich, 2017). And MAC suggests that these seven first-order 'moral elements' may combine to form 21 second-order 'moral molecules' (and 35 third-order molecules, and so on). For example, Family and Deference may combine to form Filial Piety (Nichols, 2013). Investigating how dilemmas arise and are resolved, and how higher-order concepts emerge, could extend the explanatory scope of the theory further still.

Fourth, the present study has looked for invariant aspects of morality, in two English-speaking Western, Educated, Industrialised, Rich, Democratic cultures (W.E.I.R.D; Henrich, Heine, & Norenzayan, 2010). Future research should test this factor-structure in a wider range of languages and cultures, and investigate how morality varies between individuals, within and between cultures – especially, whether moral values reflect the value of different types of cooperation under different social conditions.

Fifth, reviewing the final set of items, we note the possibility that some may be less-than-optimally phrased. For example, it's possible that questions that use comparative ("there should be *more* X"), superlative ("Y is the *most* admirable trait") or extreme ("you should *always* do Z") terms may be somewhat ambiguous, and hence difficult to interpret. Participants may value X, but disagree with the item because they think there is enough of it; they may admire Y, but disagree with because they think it is the second most important trait; or they may endorse X, but disagree because they can conceive of plausible exceptions. This applies to MAC-Q items like "Society should do more to honour its heroes", "Courage in the face of adversity is the most admirable trait", and "You should always be loyal to your family", as well as MFQ items like "It is more important to be a team player than to express oneself", "Compassion for those who are suffering is the most crucial virtue", "It can never be right to kill a human being". Future research should experiment with simpler positive language, being mindful of the ceiling and floor effects that 'milder' items may produce. We also note that the MAC-Q Heroism item "To be willing to lay down your life for your country is the height of bravery" introduces a possible confound with Mutualism, and should be avoided in future. In addition, we note that the Division items focus on the simplest form of fairness: equality. Future research should aim to explore other more nuanced expressions of fairness, such as proportionality or merit, which can lead to unequal outcomes (Starmans, Sheskin, & Bloom, 2017). Lastly, the item selection procedure delivered reversed Judgement items for Property, and for Property only, which may have introduced a confound in the valence of the items. Future work should investigate this, and if necessary correct it. Generally speaking, future research should aim to replicate the present findings with alternative sets of items, and indeed with other types of stimuli (such as standardised vignettes) (Clifford, Iyengar, Cabeza, & Sinnott-Armstrong, 2015).

Finally, the present study found that the MAC-Q's psychometrics performed well, and compare favourably to the MFQ's, but there is room for improvement, especially with regard to external criterion scales. Future research should aim to identify external scales which ask questions more directly related to the moral valence of the behaviour (rather than, as noted above, asking whether a person performs that behaviour). Such research should also extend beyond self-report scales to use performances on tasks, and behavioural measures such as experimental games.

9. Conclusion

Here we have introduced the theory of Morality-as-Cooperation, and shown how it provides a principled, predictive and productive approach to the content and structure of morality. Using cooperation as our compass, we have charted a new course, and drawn up a more accurate map of the moral landscape – revealing familiar ground in greater detail, and surveying previously unexplored territory. Thus equipped, with map *and* compass, we look forward to further discoveries ahead.

Appendix A

Morality-as-Cooperation Questionnaire: Relevance Items
When you decide whether something is right or wrong, to what extent are the following considerations relevant to your thinking? (*0–100; not at all relevant, not very relevant, slightly relevant, somewhat relevant, very relevant, extremely relevant*)

Family
  Whether or not someone acted to protect their family.
  Whether or not someone helped a member of their family.
  Whether or not someone's action showed love for their family.

Group
  Whether or not someone acted in a way that helped their community.
  Whether or not someone helped a member of their community.
  Whether or not someone worked to unite a community.

Reciprocity
  Whether or not someone did what they had agreed to do.
  Whether or not someone kept their promise.
  Whether or not someone proved that they could be trusted.

Heroism
  Whether or not someone acted heroically.
  Whether or not someone showed courage in the face of adversity.
  Whether or not someone was brave.

Deference
  Whether or not someone deferred to those in authority.
  Whether or not someone disobeyed orders.
  Whether or not someone showed respect for authority.

Fairness
  Whether or not someone kept the best part for themselves.
  Whether or not someone showed favouritism.
  Whether or not someone took more than others.

Property
  Whether or not someone vandalised another person's property.
  Whether or not someone kept something that didn't belong to them.
  Whether or not someone's property was damaged.

---

[15] For correlations between the MAC-Q and MFQ subscales, see Table S33. For correlations between the MAC-Q domains and the Big 5 personality traits, see Table S34 (Rammstedt & John, 2007). And for an initial outline of how cooperative theory might apply sexual morality, see (Curry et al., 2019: Reply to commentaries).

Appendix B

Morality-as-Cooperation Questionnaire: Judgement Items

To what extent do you agree with the following statements?
(0–100; strongly disagree, disagree, neither agree or disagree, agree, strongly agree)

Family

People should be willing to do anything to help a member of their family.

You should always be loyal to your family.

You should always put the interests of your family first.

Group

People have an obligation to help members of their community.

It's important for individuals to play an active role in their communities.

You should try to be a useful member of society.

Reciprocity

You have an obligation to help those who have helped you.

You should always make amends for the things you have done wrong.

You should always return a favour if you can.

Heroism

Courage in the face of adversity is the most admirable trait.

Society should do more to honour its heroes.

To be willing to lay down your life for your country is the height of bravery.

Deference

People should always defer to their superiors.

Society would be better if people were more obedient to authority.

You should respect people who are older than you.

Fairness

Everyone should be treated the same.

Everyone's rights are equally important.

The current levels of inequality in society are unfair.

Property

It's acceptable to steal food if you are starving. (R)

It's ok to keep valuable items that you find, rather than try to locate the rightful owner. (R)

Sometimes you are entitled to take things you need from other people. (R)

Note: (R) = reverse coded.

Appendix C. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jrp.2018.10.008.

References

Alvard, M. (2001). Mutualistic hunting. In C. Stanford & H. Bunn (Eds.), *Meat-eating and human evolution* (pp. 261–278). New York: Oxford University Press.

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.

Baier, A. (1995). *Moral prejudices: Essays on ethics*. Harvard University Press.

Balliet, D., Wu, J., & De Dreu, C. K. W. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin, 140*(6), 1556–1581. https://doi.org/10.1037/a0037737.

Barclay, P. (2016). Biological markets and the effects of partner choice on cooperation and friendship. *Current Opinion in Psychology, 7*, 33–38. https://doi.org/10.1016/j.copsyc.2015.07.012.

Becker, L. C. (1977). *Property rights: Philosophic foundations*. London, Henley and Boston: Routledge & Kegan Paul.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588–606. https://doi.org/10.1037/0033-2909.88.3.588.

Berlin, I. (1997). *Two concepts of liberty the proper study of mankind: An anthology of essays*. London: Pimlico.

Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature, 442*(7105), 912–915. https://doi.org/10.1038/Nature04981.

Binmore, K. (1994a). *Game theory and the social contract, Vol 1: Playing Fair?* (Vol I) Cambridge, Massachusetts: MIT Press.

Binmore, K. (1994b). *Game theory and the social contract, Vol 2: Just playing* (Vol. II) Cambridge, Massachusetts: MIT Press.

Bissonnette, A., Perry, S., Barrett, L., Mitani, J. C., Flinn, M., Gavrilets, S., & de Waal, F. B. M. (2015). Coalitions in theory and reality: A review of pertinent variables and processes. *Behaviour, 152*(1), 1–56. https://doi.org/10.1163/1568539X-00003241.

Bobbio, A., Nencini, A., & Sarrica, M. (2011). Il Moral Foundation Questionnaire: Analisi della struttura fattoriale della versione italiana [The Moral Foundations Questionnaire: Analysis of the factor structure of the Italian version]. *Giornale di Psicologia, 5*, 7–18.

Boone, J. (1992). Competition, cooperation and the development of social hierarchies. In E. A. Smith & B. Winterhalder (Eds.), *Ecology, evolution and social behavior* (pp. 301–337). Aldine de Gruyter.

Boos, M., Kolbe, M., Kappeler, P. M., & Ellwart, T. (Eds.). (2011). *Coordination in human and primate groups*. Springer.

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *PNAS, 108*, 10918–10925. https://doi.org/10.1073/pnas.1100290108.

Brams, S. J., & Taylor, A. D. (1996). *Fair division: From cake-cutting to dispute resolution*. Cambridge University Press.

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

Brosnan, S. F., & de Waal, F. B. M. (2014). Evolution of responses to (un)fairness 1251776-1251776. *Science (New York, N.Y.), 346*(6207). https://doi.org/10.1126/science.1251776.

Byrne, B. M., & van de Vijver, F. J. R. (2014). Factorial structure of the family values scale from a multilevel-multicultural perspective. *International Journal of Testing, 14*(2), 168–192. https://doi.org/10.1080/15305058.2013.870903.

Carter, G. G. (2014). The reciprocity controversy. *Animal Behavior and Cognition, 1*(3), 368–386.

Chilton, B. D., & Neusner, J. (Eds.). (2009). *The golden rule: The ethics of reciprocity in world religions*. Bloomsbury Academic.

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.

Confucius (1994). A single word. In P. Singer (Ed.), *Ethics* (pp. 76). Oxford: Oxford University Press.

Cosmides, L., & Tooby, J. (2005). Neurocognitive adaptations designed for social exchange. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 584–627). NY: Wiley.

Cross, S. E., Uskul, A. K., Gerçek-Swing, B., Sunbay, Z., Alözkan, C., Günsoy, C., ... Karakitapoğlu-Aygün, Z. (2014). Cultural prototypes and dimensions of honor. *Personality and Social Psychology Bulletin, 40*(2), 232–249. https://doi.org/10.1177/0146167213510323.

Curry, O. S. (2005). Morality as natural history: An adaptationist account of ethics. (PhD). London: London School of Economics. Retrieved from http://etheses.lse.ac.uk/2/.

Curry, O. S. (2007). The conflict-resolution theory of virtue. In W. P. Sinnott-Armstrong (Ed.). *Moral psychology* (Vol. I, pp. 251–261). Cambridge, Massachusetts: MIT Press.

Curry, O. S. (2016). Morality as cooperation: A problem-centred approach. In T. K. Shackelford & R. D. Hansen (Eds.), *The evolution of morality* (pp. 27–51). Springer International Publishing.

Curry, O. S., & Jones Chesters, M. (2012). 'Put yourself in the other fellow's shoes': The role of 'theory of mind' in solving coordination problems. *Journal of Cognition and Culture, 12*, 147–159.

Curry, O. S., Price, M. E., & Price, J. G. (2008). Patience is a virtue: Cooperative people have lower discount rates. *Personality and Individual Differences, 44*, 778–783.

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).

Davey, L., Bobocel, D. R., Son Hing, L., & Zanna, M. (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.

Dawkins, R. (1976/2006). *The selfish gene* (3rd ed.). Oxford: Oxford University Press.

Dawkins, R. (1979). Twelve misunderstandings of kin selection. *Zeitschrift für Tierpsychologie, 51*(2), 184–200. https://doi.org/10.1111/j.1439-0310.1979.tb00682.x.

Dawkins, R. (1998). *Unweaving the rainbow: Science, delusion and the appetite for wonder*. London: Penguin Books.

DeScioli, P. (2016). The side-taking hypothesis for moral judgment. *Current Opinion in Psychology, 7*, 23–27. https://doi.org/10.1016/j.copsyc.2015.07.002.

DeScioli, P., & Karpoff, R. (2015). People's judgments about classic property law cases. *Human Nature, 26*(2), 184–209. https://doi.org/10.1007/s12110-015-9230-y.

DeScioli, P., & Kurzban, R. (2009). Mysteries of morality. *Cognition, 112*(2), 281–299. https://doi.org/10.1016/j.cognition.2009.05.008.

DeScioli, P., & Kurzban, R. (2013). A solution to the mysteries of morality. *Psychological Bulletin, 139*(2), 477–496. https://doi.org/10.1037/a0029065.

Desimone, J. A., Harms, P. D., & Desimone, A. J. (2015). Best practice recommendations for data screening. *Journal of Organizational Behavior, 36*(2), 171–181. https://doi.org/10.1002/job.1962.

Dinno, A. (2009). Exploring the sensitivity of horn's parallel analysis to the distributional form of random data. *Multivariate Behavioral Research, 44*(3), 362–388. https://doi.org/10.1080/00273170902938969.

Downie, R. S. (1965). Forgiveness. *Philosophical Quarterly, 15*(59), 128–134.

Dugatkin, L. A. (1997). *Cooperation among animals: An evolutionary perspective*. New York: Oxford University Press.

Edel, M., & Edel, A. (1959/1968). *Anthropology and ethics: The quest for moral understanding*. Cleveland, Ohio: Case Western Reserve University Press.

Eisenberger, R., Lynch, P., Aselage, J., & Rohdieck, S. (2004). Who takes the most revenge? Individual differences in negative reciprocity norm endorsement. *Personality and Social Psychology Bulletin, 30*(6), 787–799. https://doi.org/10.1177/0146167204264047.

Emmons, R. A. (Ed.) (2004). The psychology of gratitude. OUP.

Everett, J. A. C. (2013). The 12 Item Social and Economic Conservatism Scale (SECS). *PLoS ONE, 8*(12), e82131. https://doi.org/10.1371/journal.pone.0082131.

Fiddick, L., Cummins, D., Janicki, M., Lee, S., & Erlich, N. (2013). A cross-cultural study of noblesse oblige in economic decision-making. *Human Nature, 24*(3), 318–335. https://doi.org/10.1007/s12110-013-9169-9.

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika, 10*(4), 507–521. https://doi.org/10.2307/2331838.

Fiske, A., & Rai, T. S. (2014). *Virtuous violence: Hurting and killing to create, sustain, end, and honor social relationships*. Cambridge University Press.

Friedman, O., & Neary, K. R. (2008). Determining who owns what: Do children infer ownership from first possession? *Cognition, 107*(3), 829–849. https://doi.org/10.1016/j.cognition.2007.12.002.

Fukuyama, F. (1996). *Trust: The social virtues and the creation of prosperity*. London: Penguin Books.

Gardner, A., & West, S. A. (2014). Special issue: Inclusive fitness; 50 years on. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*.

Georgas, J., Berry, J. W., Van de Vijver, F. J., Kagitçibasi, Ç., & Poortinga, Y. H. (Eds.). . *Families across cultures*. Cambridge University Press.

Gert, B. (2013). Loyalty and morality. In S. Levinson, J. Parker, & P. Woodruff (Eds.). *Nomos* (Vol. LIV, pp. 3–21). New York & London: New York University Press.

Gibbard, A. (1990a). Norms, discussion, and ritual: Evolutionary puzzles. *Ethics, 100* (July), 787–802.

Gibbard, A. (1990b). *Wise choices, apt feelings*. Oxford: Clarendon Press.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Harvard University Press.

Gintis, H. (2007). The evolution of private property. *Journal of Economic Behavior & Organization, 64*(1), 1–16.

Gintis, H., Smith, E. A., & Bowles, S. (2001). Costly signaling and cooperation. *Journal of Theoretical Biology, 213*, 103–119. https://doi.org/10.1006/jtbi.2001.2406.

Glover, R. J., Natesan, P., Wang, J., Rohr, D., McAfee-Etheridge, L., Booker, D. D., ... Wu, M. W. (2014). Moral rationality and intuition: An exploration of relationships between the Defining Issues Test and the Moral Foundations Questionnaire. *Journal of Moral Education, 43*(4), 395–412. https://doi.org/10.1080/03057240.2014.953043.

Godfray, H. C. J. (1992). The evolution of forgiveness. *Nature, 355*, 206–207.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. C. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*, 84–96.

Gómez, Á., Brooks, M. L., Buhrmester, M. D., Vázquez, A., Jetten, J., & Swann, W. B. Jr, (2011). On the nature of identity fusion: Insights into the construct and a new measure. *Journal of Personality and Social PSychology, 100*(5), 918–933. https://doi.org/10.1037/a0022642.

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. In D. Patricia & P. Ashby (Eds.). *Advances in experimental social psychology* (Vol. 47, pp. 55–130). Academic 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., 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. (2015). The rise of moral cognition. *Cognition, 135*, 39–42. https://doi.org/10.1016/j.cognition.2014.11.018.

Haidt, J. (2007). The new synthesis in moral psychology. *Science, 316*, 998–1002.

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research, 20*(1).

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

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.

Haidt, J., & Kesebir, S. (2010). Morality. In S. Fiske, G. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed., pp. 797–832). Hoboken, NJ: Wiley.

Hamilton, W. D. (1963). The evolution of altruistic behavior. *The American Naturalist, 97*, 354–356.

Hammerstein, P. (Ed.). (2003). *Genetic and cultural evolution of cooperation*. Cambridge, Mass: MIT Press.

Harcourt, A., & de Waal, F. B. M. (Eds.). (1992). *Coalitions and alliances in humans and other animals*. Oxford: Oxford University Press.

Hardy, C. W., & Briffa, M. (Eds.). (2013). *Animal contests*. Cambridge: Cambridge University Press.

Hare, D., Reeve, H. K., & Blossey, B. (2016). Evolutionary routes to stable ownership. *Journal of Evolutionary Biology*. https://doi.org/10.1111/jeb.12859.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., ... Tracer, D. (2005). 'Economic Man' in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences, 28*(6), 795–855.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Beyond WEIRD: Towards a broad-based behavioral science. *Behavioral and Brain Sciences, 33*(2–3), 111–135. https://doi.org/10.1017/S0140525x10000725.

Herskovits, M. J. (1952). *Economic anthropology: A study in comparative economics*. New York: Afred A. Knopf.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*(2), 179–185. https://doi.org/10.1007/bf02289447.

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

Huntingford, F. A., & Turner, A. K. (1987). *Animal conflict*. London & New York: Chapman and Hall.

Inglehart, R., & Baker, W. E. (2000). Modernization, cultural change, and the persistence of traditional values. *American Sociological Review, 65*(1), 19–51. https://doi.org/10.2307/2657288.

Iurino, K., & Saucier, G. (2018). Testing the measurement invariance of the moral foundations questionnaire across 27 countries (unpublished manuscript).

Jaeggi, A. V., & Gurven, M. (2013). Reciprocity explains food sharing in humans and other primates independent of kin selection and tolerated scrounging: A phylogenetic meta-analysis. *Proceedings of the Royal Society of London B: Biological Sciences, 280*(1768).

Johnson, D. D. P., & Toft, M. D. (2014). Grounds for war: The evolution of territorial conflict. *International Security, 38*(3), 7–38. https://doi.org/10.1162/ISEC_a_00149.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 263–291.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guildford Press.

Krasnow, M. M., Delton, A. W., Cosmides, L., & Tooby, J. (2016). Looking under the hood of third-party punishment reveals design for personal benefit. *Psychological Science, 27*(3), 405–418. https://doi.org/10.1177/0956797615624469.

Kurland, J., & Gaulin, S. (2005). Cooperation and conflict among kin. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 447–482). Hoboken, New Jersey: John Wiley & Sons.

Lehmann, L., & Keller, L. (2006). The evolution of cooperation and altruism – A general framework and a classification of models. *Journal of Evolutionary Biology, 19*(5), 1365–1376. https://doi.org/10.1111/j.1420-9101.2006.01119.x.

Lewis, D. K. (1969). *Convention: A philosophical study*. Cambridge, Massachusetts: Harvard University Press.

Lewis, G. J., & Bates, T. C. (2010). Genetic evidence for multiple biological mechanisms underlying in-group favoritism. *Psychological Science, 21*(11), 1623–1628. https://doi.org/10.1177/0956797610387439.

Lieberman, D., Tooby, J., & Cosmides, L. (2003). Does morality have a biological basis? An empirical test of the factors governing moral sentiments relating to incest. *Proceedings of the Royal Society of London Series B-Biological Sciences, 270* (1517), 819–826. https://doi.org/10.1098/Rspb.2002.2290.

Lieberman, D., Tooby, J., & Cosmides, L. (2007). The architecture of human kin detection. *Nature, 445*(7129), 727–731. https://doi.org/10.1038/Nature05510.

Locke, J. (2000). Labour as the basis of property. In M. Rosen, W. Wolff, & C. McKinnon (Eds.), *political thought* (pp. #73). Oxford: Oxford University Press.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130–149.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84–99.

MacIntyre, A. C. (1981a). *After virtue*. London: Gerald Duckworth & Co., Ltd.

MacIntyre, A. C. (1981b). The nature of the virtues. *Hastings Center Report, 11*(2), 27–34.

Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge: Cambridge University Press.

Maynard Smith, J., & Price, G. R. (1973). The logic of animal conflict. *Nature, 246*, 15–18.

Maynard Smith, J., & Szathmáry, E. (1995). *The major transitions in evolution*. Oxford: Oxford University Press.

Mazur, A. (2005). *Biosociology of dominance and deference*. Lanham, Maryland: Rowan & Littlefield.

McElreath, R., Boyd, R., & Richerson, P. J. (2003). Shared norms and the evolution of ethnic markers. *Current Anthropology, 44*(1), 122–129.

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17*(3), 437–455. https://doi.org/10.1037/a0028085.

Messick, D. M. (1993). Equality as a decision rule. In B. Mellers & J. Baron (Eds.), *Psychological perspectives on justice* (pp. 11–31). Cambridge: Cambridge University Press.

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus User's Guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Muthukrishna, M., Francois, P., Pourahmadi, S., & Henrich, J. (2017). Corrupting cooperation and how anti-corruption strategies may backfire. *Nature Human Behaviour, 1,* 0138. https://doi.org/10.1038/s41562-017-0138 https://www.nature.com/articles/s41562-017-0138#supplementary-information.

Nagel, T. (1991). Mortal questions: Canto.

Nash, J. (1950). The bargaining problem. *Econometrica,* 155–162.

Nichols, R. (2013). The origins and effects of filial piety (Xiao 孝): How culture solves an evolutionary problem for parents. *Journal of Cognition and Culture, 13,* 201–230.

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.

Nunn, C. L., & Lewis, R. J. (2001). Cooperation and collective action in animal behaviour. In *Economics in nature: Social Dilemmas, mate choice and biological markets* (pp. 42–66). Cambridge: CUP.

Ohtsubo, Y., & Watanabe, E. (2009). Do sincere apologies need to be costly? Test of a costly signaling model of apology. *Evolution and Human Behavior, 30*(2), 114–123.

Ostrom, L., & Walker, J. (Eds.). (2002). *Trust and reciprocity: Interdisciplinary lessons from experimental research.* New York: Russell Sage Foundation.

Park, N., Peterson, C., & Seligman, M. E. P. (2006). Character strengths in fifty-four nations and the fifty US states. *The Journal of Positive Psychology, 1*(3), 118–129. https://doi.org/10.1080/17439760600619567.

Pennock, J. R., & Chapman, J. W. (1979). *Compromise in ethics, law and politics* (Vol. XXI). New York: New York University Press.

Pennock, J. R., & Chapman, J. W. (1980). *Property* (Vol. XXII). New York: New York University Press.

Perugini, M., Gallucci, M., Presaghi, F., & Ercolani, A. P. (2003). The personal norm of reciprocity. *European Journal of Personality, 17*(4), 251–283. https://doi.org/10.1002/per.474.

Petersen, M. B. (2013). Moralization as protection against exploitation: Do individuals without allies moralize more? *Evolution And Human Behavior, 34*(2), 78–85. https://doi.org/10.1016/j.evolhumbehav.2012.09.006.

Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification.* Washington, DC/New York: American Psychological Association/Oxford University Press.

Pinker, S. (2010). The cognitive niche: Coevolution of intelligence, sociality, and language. *Proceedings of the National Academy of Sciences of the United States of America, 107,* 8993–8999. https://doi.org/10.1073/pnas.0914630107.

Popper, K. R. (1945). *The open society and its enemies.* London: Routledge.

Preuschoft, S., & van Schaik, C. P. (2000). Dominance and communication: Conflict management in various social settings. In F. Aureli & F. B. M. de Waal (Eds.), *Natural conflict resolution* (pp. 77–105). Berkeley: University of California 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.

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41,* 203–212.

Rawls, J. (1958). Justice as fairness. *The Philosophical Review, 67*(2), 164–194.

Reynolds, W. M. (1982). Development of reliable and valid short forms of the marlowe-crowne social desirability scale. *Journal of Clinical Psychology, 38*(1), 119–125. https://doi.org/10.1002/1097-4679(198201)38:1<119::AID-JCLP2270380118>3.0.CO;2-I.

Richards, N. (1988). Forgiveness. *Ethics, 99*(1), 77–97.

Riechert, S. E. (1998). Game theory and animal contests. In L. A. Dugatkin & H. K. Reeve (Eds.), *game theory and animal behavior* (pp. 64–93). Oxford: Oxford University Press.

Robinson, D. R., & Goforth, D. J. (2005). *The topology of 2x2 games: A new periodic table.* London: Routledge.

Rose, C. M. (1985). Possession as the origin of property. *University of Chicago Law Review,* 73–88.

Royce, J. (1908). *The philosophy of loyalty.* New York: Macmillan.

Rubin, P. H. (2000). Hierarchy. *Human Nature, 11*(3), 259–279.

Sachs, J. L., Mueller, U. G., Wilcox, T. P., & Bull, J. J. (2004). The evolution of cooperation. *The Quarterly Review of Biology, 79*(2), 135–160. https://doi.org/10.1086/383541.

Sartre, J. P. (1946/1973). Existentialism and humanism. London: Methuen.

Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H. Heijmans, D. S. G. Pollock, & A. Satorra (Eds.), *Innovations in multivariate statistical analysis: A Festschrift for Heinz Neudecker* (pp. 233–247). Boston, MA: Springer, US.

Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika, 75*(2), 243–248. https://doi.org/10.1007/s11336-009-9135-y.

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: Theory and empirical tests in 20 countries. In M. Zanna (Ed.). *Advances in Experimental Social Psychology* (Vol. 25, pp. 1–65). New York: Academic Press.

Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger. *Proceedings of the National Academy of Sciences of the United States of America, 106*(35), 15073–15078. https://doi.org/10.1073/Pnas.0904312106.

Shackelford, T. K., & Hansen, R. D. (Eds.). (2016). *The evolution of morality.* Springer International Publishing.

Sherratt, T. N., & Mesterton-Gibbons, M. (2015). The evolution of respect for property. *Journal of Evolutionary Biology, 28*(6), 1185–1202. https://doi.org/10.1111/jeb.12648.

Shultz, S., Opie, C., & Atkinson, Q. D. (2011). Stepwise evolution of stable sociality in primates. *Nature, 479*(7372), 219–222 :http://www.nature.com/nature/journal/v479/n7372/abs/nature10601.html#supplementary-information.

Sinnott-Armstrong, W. (2007). *Moral psychology* (Vol. 1–3). Cambridge, Massachusetts: MIT Press.

Skyrms, B. (1996). *Evolution of the social contract.* Cambridge: Cambridge University Press.

Smith, K. B., Alford, J. R., Hibbing, J. R., Martin, N. G., & Hatemi, P. K. (2016). Intuitive ethics and political orientations: Testing moral foundations as a theory of political ideology. *American Journal of Political Science.* https://doi.org/10.1111/ajps.12255.

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

Sterelny, K., & Fraser, B. (2016). Evolution and moral realism. *British Journal for the Philosophy of Science, 68*(4), 981–1006.

Strassmann, J. E., & Queller, D. C. (2014). Privatization and property in biology. *Animal Behaviour.* https://doi.org/10.1016/j.anbehav.2014.02.011.

Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual Review of Psychology, 64*(1), 231–255. https://doi.org/10.1146/annurev-psych-113011-143812.

Tooby, J., & DeVore, I. (1987). The reconstruction of hominid behavioral evolution through strategic modeling. In W. G. Kinzey (Ed.), *The evolution of human behavior: Primate models* (pp. 183–237). Albany, New York: SUNY Press.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*(1), 35–57. https://doi.org/10.1086/406755.

Von Neumann, J., & Morgenstern, O. (1944). *The theory of games and economic behavior.* Princeton: Princeton University Press.

Weeden, J., & Kurzban, R. (2013). What predicts religiosity? A multinational analysis of reproductive and cooperative morals. *Evolution and Human Behavior, 34*(6), 440–445.

Westermarck, E. A. (1994). The horror of incest. In P. Singer (Ed.), *Ethics* (pp. 95–96). Oxford: Oxford University Press.

Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement, 9*(1), 1–26. https://doi.org/10.1177/014662168500900101.

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 73*(6), 913–934. https://doi.org/10.1177/0013164413495237.

Wong, D. (1984). *Moral relativity.* UC California Press.

Yilmaz, O., Harma, M., Bahcekapili, 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.