Universitat Autònoma de Barcelona

Neural mechanisms of willingness to fight and die for sacred values in the context of intercultural conflict

Clara Pretus Gómez

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons: http://cat.creativecommons.org/?page_id=184

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons: http://es.creativecommons.org/blog/licencias/

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license: https://creativecommons.org/licenses/?lang=en

Doctoral thesis

Programa de Doctorat en Psiquiatria i Psicologia Clínica
Departament de Psiquiatria i Medicina Legal
Universitat Autònoma de Barcelona

Neural mechanisms of willingness to fight and die for sacred values in the context of intercultural conflict

A neuroimaging study

Author

Clara Pretus Gómez

Director

Dr Óscar Vilarroya Oliver

Barcelona, September 2016

“Two things fill the mind with ever new and increasing admiration and awe,
the more often and steadily we reflect upon them:
the starry heavens above me and the moral law within me.
I do not seek or conjecture either of them
as if they were veiled obscurities or extravagances beyond the horizon of my vision;
I see them before me and connect them immediately
with the consciousness of my existence.”

Immanuel Kant
(Critique of Practical Reason, 1788)

“Well, ’tis no matter; honour pricks me on.
Yea, but how if honour prick me off when I come on? How then?
Can honour set-to a leg? No. Or an arm? No. Or take away the grief of a wound?
No.
Honour hath no skill in surgery, then? No. What is honour? A word.
What is in that word “honour”? What is that “honour”? Air.
A trim reckoning! Who hath it? He that died o’ Wednesday.
Doth he feel it? No. Doth he hear it? No. ’Tis insensible then? Yea, to the dead.
But will it not live with the living? No. Why? Detraction will not suffer it.
Therefore I’ll none of it. Honour is a mere scutcheon.
And so ends my catechism.“

William Shakespeare
(Henry IV Part I, 1597)

Table of contents

Acknowledgments

I would like to thank first of all my supervisor Óscar Vilarroya for his guidance and support throughout these years. I am particularly impressed by his ability to manage difficult situations in a way that gets to combine both goal-orientedness and humaneness. I would also like to thank the team in the Laboratorio de Imagen Médica, Hospital Gregorio Marañón (Madrid), especially Susanna Carmona, David García and Luís Marcos for their good vibes and support whenever I found myself in technical trouble, and the team in the Bernstein Center of Computational Neuroscience in the Charité-Humboldt Universität zu Berlin, especially Kai Görgen, from whom I learned a lot both in Neuroimaging analysis and the art of doubting. This work would not have been possible without the close collaboration with the Artis International Research team, who took me out of the lab walls to show me the exciting world of field research to find solutions to current global challenges, especially Nafees Hamid, Scott Atran and Richard Davis. In addition, I would like to thank the great professionals working in my department and who have helped me and inspired me in many ways, Daniel Bergé, Paula Calvo and Jaume Fatjó.

Last not least, my work should be acknowledged to my undoubtedly inspiring parents, Àuria and Joan, who have bred my curiosity for natural phenomena from a very early age and who have transmitted me the urge to get a deeper knowledge of human nature as a necessary step to improve our relationship with ourselves, our fellows and the environment. I would also like to mention here my little crowd of “sacred” people, Alexandra, Candela, Marta, David and Xavi, with whom I have grown into who I am now, and who have shown me the importance of caring for each other and believing in ourselves.

Thanks to all of you!
Moltes gràcies!

Key concepts

devoted actors

are those decision-makers who are willing to engage in costly sacrifices irrespective of risks and prospects of success in defense of their moral convictions, thus defying rationalistic accounts of decision-making typically based on cost-benefit analysis (Atran, 2006). A more detailed explanation of the devoted actor model of decision making follows in section 2.3.

inter-group conflict

or identity-based conflict arises when members of social groups behold a collective identity that is essentially different than that of their rivals, tracing a defining line between “them” and “us” typically based on race, religion or gender, thus eliciting antagonistic attitudes towards opponents (Kriesberg, 2003).

intractable conflict

is defined as a long-term dispute that is seemingly impossible to resolve given the level of escalation and destructiveness attained. It is usually caused and perpetuated by deep-rooted differences in moral foundations stemming from different cultural viewpoints (with all involved parties defending what they see as fundamentally “right”), inequitable distribution of highly valuable goods such as land, water or wealth, and the strive for social and political dominance (Burgess & Burgess, 2003).

moral community

is employed here in reference to any group of people who are socially and morally integrated, that is, who hold intensive and extensive attachments and share a set of moral beliefs, e.g. a religious or military group (Marshall, 1998).

social identity

refers to an individual’s understanding of the self as a member of a relevant social group (Turner & Oakes, 1986) that can be used to predict intergroup attitudes and behaviors (Tajfel & Turner, 1979).

Abstract

Intercultural conflict is often propagated and sustained through an avid defense of opposing cultural views, both religious and nationalistic, that moral community members are ready to defend even at the cost of their own lives. Such group-defining *sacred values* or extremely overvalued beliefs are only poorly understood in terms of their neurobiological substrate, with no neuroimaging studies conducted in the context of real intercultural conflict. This is the first neuroimaging study aimed to identify the neural mechanisms underlying willingness to fight and die (WFD) in defense of sacred values (SVs). Moreover, it is the first to assess sacred values in a real-world sample of Muslim Pakistani males recruited in the area of Barcelona (Spain), who expressed radicalized political views with respect to the Kashmiri conflict. A functional magnetic resonance paradigm was employed to evaluate both the neural activity associated with WFD for sacred values and the effect of social influence on the same measure, thus assessing post-manipulation moral outrage reactions and possible induced differences in WFD ratings. As a result, WFD for SVs elicited neural activity in areas associated with deontic thinking, affective appraisal of self-relevant content and self-agency (vmPFC/vACC), and decreased activity in neural nodes associated with utilitarian reasoning, deliberation and introspection (dmPFC, dlPFC, IFG). In turn, conflicting social feedback elicited moral outrage responses, which predicted activity in the parietal operculum, associated with pain and disgust. However, WFD ratings were sensitive to social influence, changing in the direction established by the norm. These results suggest that WFD for SVs is primarily driven by affective-laden self-enhancing motives that presumably respond to social expectations, pointing at acts of extreme altruism as social validation tools conferring feelings of personal empowerment. Support for different de-radicalization approaches is discussed in the end.

1. Introduction

Ever since the New York twin tower attacks in 2001, large-scale terrorism has increasingly become a first-line worldwide concern that is still on the wait for efficient solutions, leaving thousands of civilian victims behind every year. From highly organized big-impact actions such as the Paris attacks on November 13th 2015 up to domestic lower scale incidents such as the Würzburg train attacks by a 17 year old on July 2016, numerous violent symbolic actions against innocents have been taking place around the planet in the last few years, especially hitting Muslim cities such as Istanbul, Kabul, Beirut, Cairo or Baghdad. Whereas the weapons used in the different actions are diverse, ranging from Kalashnikovs and explosives up to kitchen knives or even trucks, the apparent common motive by perpetrators seems to be sworn loyalty to the jihadist movement. However, most of the perpetrators seem to have rather little knowledge of Islam, pointing to causes external to the religion *per se*, such as internalized political attitudes. More generally, the main question revolves around the causes that motivate particular individuals to engage in costly sacrifices in defense of specific cultural views, a matter of intense debate by a myriad of specialists in different fields, from military strategists to behavioral scientists. In the present work, we present a neuroimaging perspective on the topic, testing individuals with radicalized views on a specific intergroup conflict on their motivation to fight and die in defense of their personal cultural values in an fMRI scanner.

1. 1. What are sacred values?

Tensions underlying intercultural conflict often arise from conflicting cultural views that become highly protected or *sacralized* by the members of a moral community (Burgess & Burgess, 2003). In the present work, the term *sacred values* (SVs) will be employed to refer to those values that are considered to have an absolute significance by individuals belonging to a given social group, without any necessary reference to religious content. Such absolutist views cannot be measured in economic or materialistic terms, thereby escaping cost-benefit analysis (Baron & Spranca, 1997). In other words, SVs are treated as premises that should always hold true, no matter what other circumstances stand on the way. Although it may seem a delicate concept to pinpoint, SVs count on extensive

previous psychological and anthropological research on the field (see section 2.2). Whereas some values appear to be a constant motto through human societies, e.g. acting with fairness (Rochat, Dias, Broesch, Passos-Ferreira, & Berg, 2009), others might only hold true for a subset of them, e.g. voting democratically or preserving a woman's virginity as a moral value. Hence, opposing interpretations of a given value by different moral communities, especially if the value is hold sacred, can derive into serious intercultural conflict (Burgess & Burgess, 2003). Thereby, the study of sacred values appears as a main pillar providing further insight into the nature of identity-based conflicts.

As shown in multiple studies (Baron & Spranca, 1997; McGraw & Tetlock, 2005; Tetlock, 2003), people refuse to give up their personal sacred values in exchange for material benefits, what is known as *taboo trade-offs* (Duc, Hanselmann, Boesiger, & Tanner, 2013; Hanselmann & Tanner, 2008). As a prime example of this, people are not willing to stop believing in god in exchange for any amount of money (Berns et al., 2012). Indeed, instrumentalization and violation of sacred values is associated with intense emotional reactions of disgust and moral outrage (McGraw & Tetlock, 2005; Tetlock, 2003). Given their inherent innegotiability, sacred values establish boundaries to people's life choices and drive action in ways that cannot be predicted by rational actor models of decision-making, e.g. prohibiting abortion regardless of the circumstances because "life is sacred". Consequentialist models only account for "cold" outcome-oriented cost-benefit calculations irrespective of the values at stake (see Box 1 for greater detail). Herein, they fail to predict behavior governed by sacred values, which rather responds to deontic reasoning styles involving absolute "goods" and "bads". Hence, new decision-making models focus on the striking role of sacred values, moral beliefs and no-go situations that sculpt people's choices within what are known as devoted actor models of decision-making (Scott Atran, 2003) (see section 2.3).

Importantly, people's social identity and group membership is often bound to a particular set of sacred values (Sachdeva & Medin, 2009; Swann, Jetten, Gómez, Whitehouse, & Bastian, 2012), so much so that violation of group-specific sacred values may be punished by law and even lead to social exclusion of the perpetrator, e.g. if he does not comply with the established norms. Likewise, and as introduced above, conflicting sacred values between different moral communities may lead to inter-group conflict, which throughout

time may further create new sacred values in relation to the out-group and the conflict itself. Thus, inter-group tensions involving sacred values can quickly escalate into deep-seated intractable conflicts that persist across generations (see section 2.2.2. for specific examples). For these reasons, and given their power to create strong commitment, awaken hostility and mobilize action, the study of sacred values in the context of intercultural conflict seems highly relevant, especially in the face of the last episodes of terrorism that have shaken Europe and the rest of the world. In the best possible scenario, research on sacred values might give important insights into the nature of long-term inter-cultural conflicts, as well as shed light on realistic ways to negotiation and peacemaking.

Box 1. Consequentialist/utilitarian vs. Deontologist moral philosophy

Consequentialist and deontological moral theories have created a major divide within the Western philosophy that has been extended for centuries. On one hand, the utilitarian or consequentialist tradition defended by authors such as Stuart Mill (1806-1873) postulates that choices can be morally assessed as a function of the amount of "Good" (or valuable states of affairs) they produce. Thereby, any action that promotes a greater good is right, thus prioritizing the good over the right. Nonetheless, such focus on the final outcome of choices makes any kind of act justifiable regardless of how harmful it is. For instance, it permits that innocents are killed or mistreated as long as these actions generate a greater benefit for others (Alexander & Moore, 2015).

On the other side, deontologist theories, with Immanuel Kant (1724-1804) as a their main representative, establish that certain choices are required, forbidden or permitted independently of the amount of good that could otherwise be generated. That is, conformity with a moral norm by a moral agent makes a choice right regardless of its effects, prioritizing the right over the good. In this context, moral agents are expected to simply obey the moral norm, for instance, "not to kill innocents" (Alexander & Moore, 2015).

A way to assess these two seemingly opposing philosophical stances in psychology has been by means of the so-called moral dilemmas that position the subject between two behavioral options that pitch different outcomes against moral norms typically involving

"harming others". For instance, the *trolley* dilemma (Thomson, 1985) presents a scenario where a runaway trolley is about to kill five people and can only be stopped if the subject hits a switch that will deviate the trolley so that instead of killing five, it will only kill one person (see Figure 1, left). As has been observed, most participants would hit the switch (Greene et al., 2001), a choice that has been associated with utilitarian decision-making, as it responds to a call for the greater good. On the other side, in the *footbridge* dilemma (Thomson, 1985) the subjects needs to choose between letting five people get killed by a runaway trolley or pushing a stranger off a bridge in order to stop the trolley (see Figure 1, right). In this case, most participants would not push the stranger (Greene et al., 2001), which has been associated with deontic decision-making, i.e. preserving the norm ("not harming others") at the expense of the greater good.

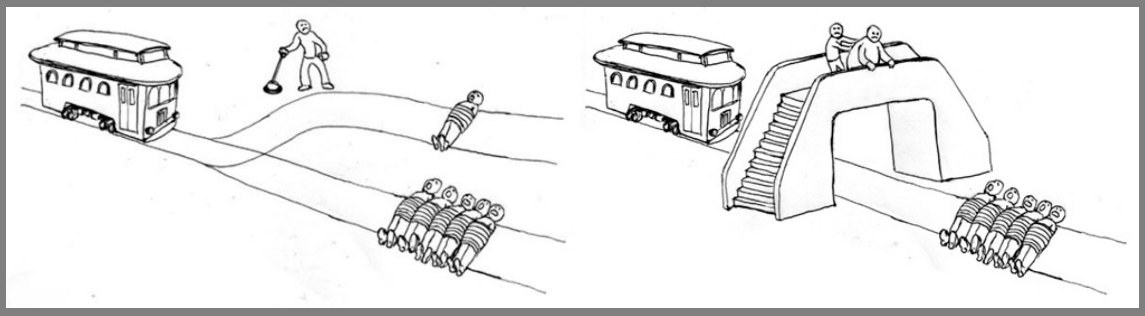


Figure 1. On the left, diagram of the trolley dilemma, where the subject can avoid the killing of five people by pushing a switch, which will yield the killing of one. On the right, diagram of the footbridge dilemma, where the subject can avoid the killing of five by pushing a person down the bridge.

1. 2. Psychological research on sacred values

1. 2. 1. Research on SVs in experimental settings

The first research around the subject of absolute values in the field of psychology started in the late nineties in two different groups. On one hand, Baron and Spranca (1997) started investigating in the line of what they called protected values (PVs), since such values, by definition, were protected from trade-offs with other values. Their work focused on biases from expected outcomes if decisions were made following utilitarian reasoning, characterized by intending to maximize the utility or amount of goodness of the outcomes (Broome, 1991). Baron's team mainly investigated sacred values related to environmentalism, such as the conservation of animal species and forests. In their studies,

people defending environmentalist PVs also exhibited an omission bias (Baron & Ritov, 2009), that is, favoring choices that involved not to intervene on the default conditions, e.g. against utilitarian reasoning, conservationists preferred 20 fish species to go extinct "naturally" rather than actively exterminating a smaller number of them in order to save those 20 species. These observations led Baron to conclude that PVs were unreflexive labile overgeneralizations that could easily be amended within a "kind of cognitive therapy", where participants would be confronted with situations and counterexamples where a small concession in a PV would involve much greater benefits (Baron & Leshner, 2000).

On the other hand, Tetlock and colleagues (2000), who coined the term Sacred Values in 1996, developed the Sacred Value Protection Model (SVPM). In his experiments, participants were exposed to *routine* trade-offs dealing with mundane values (e.g. "paying someone to clean their house") and what they called *taboo* trade-offs involving mundane against sacred values such as voting democratically. As predicted, *taboo* trade-offs (e.g. "buying and selling votes in political elections") evoked strong emotional reactions of anger and disgust in participants, norm-enforcement responses such as willingness to punish SV violators and those who failed to condemn them, and moral cleansing reactions such as willingness to volunteer in initiatives counteracting SV violations (Tetlock, Kristel, Elson, Green, & Lerner, 2000). Along these lines, Duc et al. (2013) found that situations involving two sacred values pitted against each other (e.g. "a human life against another human life") within *tragic* trade-off scenarios were perceived as particularly stressful and difficult by participants, although not outraging.

According to Fiske & Tetlock (1997), choice problems involving sacred values need to be considered within a social-relational framework, were personal decisions respond to social expectations associated with inter-individual or individual-to-society relationships. Thus, individuals' judgments do not necessarily respond to efforts to maximize utility, rather they function as identity ratification elements constrained by social roles. Therefore, sacred values can be understood as socially expected default personal choices. Such decisions define and assert an individual's social identity and, thus, need to be compatible with a coherent image of the self (McGraw & Tetlock, 2005). In line with this work, relevant aspects involved in sacred value transactions relate to which image does each of the parts involved want to transmit (e.g. the proposer and the responder). Critically, an

offer proposal might be taken as an offense if either of the parts detects a misconception in the other's views on the values articulating the social relationships at stake (McGraw & Tetlock, 2005; Tetlock et al., 2000). Therefore, in order to avoid taboo situations that could increase tensions between the two parts, it is crucial to identify which specific values are sacred for the counterpart.

1. 2. 2. Field research in areas with inter-group conflict over SVs

As presented above, experimental designs in lab settings have shown that SVs require a special treatment in trade-off situations. This is especially true for negotiations dealing with subtle subjective meanings that can easily break into strong negative emotions and the abandonment of a dialoguing stance. Therefore, in contexts of real-world inter-group conflict a deeper knowledge on cultural sacred values of each part is crucial for understanding the nature of the conflict. Such has been the goal of Atran, Ginges, Sheikh and colleagues in their body of research on SVs comprising different populations currently confronted over community goods across the globe. A paradigmatical example of such complex inter-group clashes comprising both national and religious values is the Israeli-Palestinian conflict, started in the mid-20th century and, for many, the world's most "intractable" conflict (Rice, 2014). Relevant SVs in this conflict are to establish the holy land of Israel responding to Israeli interests, and the Palestinian refugees' right to return to their homes in the colonized areas for the Arab counterpart.

Ginges, Atran, Medin, & Shikaki (2007) organized a series of experiments in the field aimed to capture inter-group attitudes and responses to realistic hypothetical peace deals involving a two-state solution in both Israeli settlers and Palestinian refugees. As expected, material sweeteners such as annual 1billion dollar payments to Israel by the US (*taboo* trade-off) elicited a *backfire* effect involving hostile reactions of anger and disgust when SVs were invoked. Conversely, symbolic incentives such as Palestinians having to renounce to their right to return (*tragic* trade-off) decreased moral outrage and support for violence. Independently from material transactions, equitable loses in the sacred domain by both sides within *tragic* trade-off scenarios decreased rejection to peace deals. Interestingly, sincere apologies by the Israeli government for past abuses to Palestinians had a remarkably positive effect over willingness to compromise sacred values in the

Palestinian community. For the political leaders who were interviewed, symbolic gestures were a necessary first step to open negotiations over material outcomes.

Further field studies addressed the Muslim-Hindu conflict in northeastern India (Sachdeva & Medin, 2009) involving the Babri Mosque and the Kashmir. The first issue is more salient for the Hindu community and concerns the sacred right over land under the Babri Mosque, claimed by Hindus and Muslims as the birthplace of one of their respective sacred figures, provoking thousands of deaths during an episode of riots and terrorist bombings in the early '90. The dispute over the Kashmir involves the struggle for Muslim autonomy within a territory split between India, Pakistan and China, presented in further detail in Box 2 as part of the present study. Similarly to Ginges et al. (2007), peace deals involving SVs and economic as well as symbolic concessions were proposed to members from both communities. In this case, whereas Hindu participants showed higher rejection to *taboo* trade-offs than Muslims in the Babri Mosque issue, presence of SVs in Muslim participants predicted disapproval of any trade-off in the Kashmir scenario, demonstrating that deal approval was subject to the salience of the conflict in the participant's social identity.

Along the same research line, Dehghani, Iliev, Atran, Ginges, & Medin (2009) investigated emerging sacred values concerning the nuclear energy program in Iran, a relatively new subject compared to the more historically rooted conflicts studied in the SV literature. The deals presented to Iranian participants involved both Iran and Israel giving up their nuclear energy program, and included either economic incentives or sanctions. The right for nuclear energy appeared to be sacred in part of the sample, which showed higher rejection rates to the *taboo* trade-offs. The emergence of new sacred values is associated with an increased perception of threat to the integrity of a given value together with the use of religious rhetoric and a narrative involving national rights and historical injustices (Dehghani et al., 2010), as discussed in the following section.

1. 2. 3. Value sacralization through perceived threat and ritualization

As seen in the Iranian example, some preferences can gain status as sacred values relatively quickly. Atran & Axelrod (2008) proposed that value sacralization may occur as a defensive response to perceived threat towards given moral community values by

competing out-groups. Indeed, a country's religiosity was found to increase as a function of existential insecurity in a cross-cultural study (Norris & Inglehart, 2004), while being reminded of death or danger exacerbated belief in god and credibility of counterintuitive happenings (Kay et al., 2010). In another study, participants primed with uncertainty showed higher resistance to accept randomness, finding more meanings and coherent patterns in an arrangement of dots (Whitson & Galinsky, 2008). Thus, whereas perceived threat and uncertainty intensifies the adherence to meaningful values, inclusion of those threatened values within religious rituals and rhetoric, boosts people's sense of sacredness with respect to them (Dehghani et al., 2009). In a series of experiments with American and Palestinian samples, Sheikh, Ginges, & Atran (2013) pinpointed the effect religious rituals and threat on value sacralization. On one hand, participation in religious ritual measured by frequency of prayer and church/mosque attendance was associated with higher number of SVs, including those which were not explicitly religious (e.g. *global warming is real/not real, North Korea should/should not be nuked*). Furthermore, priming half of the participants by asking them about the importance of religious rituals and beliefs in their lives increased the number of claimed SVs with respect to non-primed participants. On the other hand, the interaction between religiosity and probability to sacralize Palestinian "sovereignty over East Jerusalem" and "right of return" was enhanced by perceived threat to Palestinian people.

Ritualistic activity has been associated with higher commitment to group-relevant core values and in-group prosociality in several studies (Atran & Henrich, 2010). For instance, Sosis and Ruffle (2003) found an association between participating in collective ritual and in-group cooperation in Israeli kibbutzim, while costlier rituals lead to longer group survival in a study comprising over 80 utopian 19th century communes (Sosis & Bressler, 2003). However, in a field study comparing costly and collective ritual modalities (Xygalatas et al., 2013), participants who had either experienced or witnessed rituals involving higher levels of physical pain made higher donations to charity and self-identified with more inclusive groups (national identity comprising different religions), whereas participants of collective singing and prayer reported more narrow parochial identities (Hindu identity). Therefore, costly and collective rituals evoke different in-group identities and attitudes towards the out-group.

On one hand, rituals involving extreme pain have been associated with altered states of consciousness involving decreased cognitive inhibition and ability to make self-other distinctions (Lee et al., 2016). At the same time, such costly displays have an intense social impact, serving as diagnostic actions that prove the displayer's degree of commitment to his verbally stated beliefs (Henrich, 2009). Collective rituals, on the other hand, promote a sense of closed community and in-group solidarity through synchronized activity, e.g. by singing, dancing or marching (Wiltermuth & Heath, 2009). As suggested by Durkheim (1912), synchronized behavior is thought to increase group conformity in the moral domain. Indeed, Fischer (2013) found that behavioral synchrony was correlated with sacred values, a relationship that was mediated by feelings of oneness ('entitativity') and that ultimately led to higher cooperation. According to the shared goal perception-action hypothesis, synchronized action would promote acting as a unit towards shared collective goals, such as responding to a threat (Fischer, 2013; Reddish, Fischer, & Bulbulia, 2013).

In line with these observations, Ginges, Hansen, & Norenzayan (2009) found that coalitional aspects of religion such as temple attendance, but not private prayer to God, were significant predictors of support for extreme forms of parochial altruism such as suicide bombers in a Palestinian and an Israeli sample. Interestingly, primes of God as an omnipotent watcher who punishes norm violators has been related to self-monitoring of morally-relevant behavior (Norenzayan et al., 2014; Bering & Johnson, 2005) and increased public self-consciousness (Gervais & Norenzayan, 2012), presumably promoting cooperation in large groups of complex societies (Atran & Henrich, 2010). However, belief in God in itself has not been related to out-group aversion, rather it boosts pursuit of more universal goals that are inclusive of the out-groups, even among non-believers (Preston & Ritter, 2013).

In sum, whereas costly displays may work as sort of guarantee, giving credibility to the ideological commitment of the displayers (Henrich, 2009), collective rituals seem to develop the more affiliative and relational aspects of religion, including defensive attitudes and hostility towards the out-group. Although in different ways, both ritual modalities serve as a means to strengthen commitment to in-group values, which in contexts of in-group threat can turn into out-group animosity. Indeed, greater competitiveness between groups and, hence, threat to group-specific values has been associated with group rituals

that involve costlier sacrifices (Sosis, Kress, & Boster, 2007). From an evolutionary perspective, internalization of SVs through ritualistic means, especially those under threat, might have a binding effect among community members by increasing trust and in-group commitment, thus improving the group's competitiveness in the face of antagonistic conditions (Scott Atran, Axelrod, & Davis, 2007).

1. 2. 4. Pathways to value de-sacralization

On the opposite end, some researchers have dealt with ways to *unfreeze* rigid socio-cultural beliefs that originate from and, simultaneously, nourish inter-group conflict. For instance, Galinsky & Moskowitz (2000) observed that perspective-taking ("imagine a day in the life of this individual as if you were that person, looking at the world through his eyes and walking through the world in his shoes.") helped reduce stereotyping of the out-group both in an explicit and an implicit task, whereas stereotype suppression ("actively try to avoid using stereotypes") decreased explicit stereotyping but increased implicit stereotype accessibility. In turn, reduced explicit and implicit stereotyping elicited by perspective-taking was associated with a higher overlap between representations of the self and the out-group and decreased in-group favoritism. Additionally, information about losses intrinsic to a conflict has been associated with higher interest in possible solutions to the conflict and reassessment of political stances compared to receiving information about the advantages of a peace agreement (Gayer, Halperin, & Bar-tal, 2009). More generally, intergroup contact has been proven to reduce intergroup prejudice, as supported by Pettigrew & Tropp (2006) in a meta-analysis including 713 independent samples.

However, typical interventions in conflict situations involve new counter-information aimed to clash with deeply established societal beliefs about the out-group and the conflict, including reevaluation of the rival and future prospects. This strategy is intended to cause situations of cognitive dissonance and destabilize the subject's position. However, given that collective narratives in conflict situations work as constant reinforcers of out-group and conflict-related SVs, individuals might develop defensive strategies to protect them, such as ignoring, misinterpreting and ultimately rejecting incongruent information about the conflict (Hameiri, Porat, Bar-tal, Bieler, & Halperin, 2014).

Interestingly, in certain situations, seemingly counterintuitive interventions can lead to the desired outcomes. For instance, Swann, Pelham & Chidester (1988) used a paradoxical method to auspiciously decrease conservative attitudes towards women's roles by prompting participants to argue in favor of views consistent with their own, although more extreme (e.g. *Why do you think men always make better bosses than women?*). Based on these studies, Hameiri and colleagues (2014) developed a paradoxical thinking intervention within the framework of the Israeli-Palestinian conflict aimed to escape participants' rejection by avoiding counter-information and dissuading participants by conveying extreme absurd views on the conflict. In it, Jewish Israeli participants were presented a video clip transmitting the message "we need the conflict in order to have the strongest army in the world**",** thus encouraging participants to view extreme support for the army as unreasonable. After 7 waves of intervention, participants reported to have more doubts concerning whether Israelis were the ultimate victims of the conflict and whether Palestinians were "no partner for peace" than before the intervention. In turn, Israeli participants viewed Palestinians as less responsible for the conflict, which had a positive impact on their willingness to compromise over peaceful solutions. These effects were maintained after 1 year and were stronger in centrist and right-wing participants, who typically have greater adherence to anti-Palestinian values. Additionally, participants who went through the intervention were more likely to vote parties supporting peace in the general elections. For the present research, such precedent studies establish a research line for the development of de-sacralization and de-radicalization interventions with plausible outcomes under the MRI scanner.

1. 3. The devoted actor model of decision-making

1. 3. 1. Devoted actors enact extreme forms of altruism

As presented above, far from mere posturing or amenable unreflective responses, sacred values have shown to be central elements within socio-political conflicts around the globe, motivating SV-beholders to endorse extreme behaviors such as killing and sacrificing themselves in favor of abstract groups of non-kin such as religious or national communities (Ginges et al., 2009). Importantly, SVs conform the groundings of an individual's social identity, defining "who I am" within a group (Atran & Ginges, 2012). Atran et al. (2007) integrate such views into their devoted actor model of decision-making,

which portraits agents, whose behavior is ruled by SVs independently of risks and outcomes, thus challenging rational accounts of decision-making. A paradigmatic example of a devoted actor is a suicide bomber, an agent who gives up the totality of his biological self-interest in favor of a sociopolitical cause. Given the increasing worldwide concern over terrorist activity, a great deal of effort is currently being directed to disentangle the psychological mechanisms underlying such motivational stances.

The corpus of evidence collected to date characterizes suicide bombers as young ordinary people, typically unmarried males, with no history of psychopathology, suicidality or social dysfunction (such as lack of family or friends), neither do they report feeling "hopeless" (Atran, 2003). Contrary to what could be expected, support for terrorism has not been correlated with lower economic or education level. However, it has been positively correlated feelings of in-group pride and cohesiveness (Krueger, 2002). In fact, participation in political protests is correlated with identification with a social group, pointing to a higher salience of collective over individual socioeconomic or political conditions (Ginges et al., 2009). Nonetheless, according to Merari et al. (2010) the pathway from the use violence to martyrdom cannot be explained without the intervention of recruiting organizations and charismatic leaders, who deliberately manipulate its members to create strong bonds and existential commitment among fictive kin, often through emotionally loaded religious rituals. Such indoctrination might ultimately help devoted actors override attachment to self-interest to commit voluntary suicide bombings in favor of institutional goals, a strategy usually adopted by materially weaker groups confronted with stronger rivals.

1. 3. 2. Key psychological aspects of the devoted actor

A devoted actor's personal identity is in-dissociable from his in-group identity (Swann, Gómez, Seyle, Morales, & Huici, 2009). Hence, group-specific SVs have a striking role in defining "who I am" within the group, driving action in ways aimed to secure that identity (Sheikh, Gómez, & Atran, 2016). Hence, people who behold a SV will act as devoted actors in decisions that affect that value. As seen so far, such decisions will be rule-bound, responding to absolute "rights" and "wrongs", and will be taken independently of risks and outcomes. Moreover, devoted actors will negatively react at the perspective of value trade-off, a scenario where material incentives will backfire (Ginges et al., 2007).

Additionally, Sheikh et al. (2013) identified a series of devoted actor traits that seem key to the preservation of SVs over time. First, while people tend to conform to the in-group's opinion on mundane issues (Cialdini & Goldstein, 2004), devoted actors were less sensitive to social influence with respect to their SVs, preserving SV integrity. Second, the psychological distance to past and future SV-relevant events was reduced, securing present commitment to SVs. For instance, Palestinians who sacralized the right of return, perceived Palestinian exodus from Israel and an estimated future return as temporally closer events. Third, devoted actors were less willing to take individual opportunities to exit the conflict even in exchange for accomplishing other SVs unrelated to the group such as the pilgrimage to Mecca. Such reluctance to deceive the in-group is crucial to maintain collective action in the long-term.

Such characteristic pattern of behavior defies traditional rationalistic accounts of decision-making that respond to instrumental ponderation centered in an individual's self-interest. As Ginges & Atran (2010) argue, the most utilitarian choice for a rational decision-maker in a conflict situation would be to "free-ride" and expect others to do the costly work, while later benefitting from newly-conquered privileges, unless they were offered "selective incentives" or a special treatment after the future conquest. However, in samples of both Palestinians and Israeli settlers striving for self-enhancement was uncorrelated or negatively correlated with willingness to participate in violent acts, which was, conversely, predicted by higher priorization of collective values over personal ones (Ginges et al., 2010). Additionally, studies on Palestinians showed that requesting monetary compensation for a son's martyrdom was intensely disapproved and seen as taboo, while delaying martyrdom to avoid retailiative killing of a martyr's family was more strongly rejected than delaying martyrdom to look after a sick father. These two situations present substantial differences with respect to the amount of personal loss at stake and can only be predicted by the devoted actor model and SV theory.

Interestingly, Kacou (2012) defends that a martyr's interests can still respond to rational motives when considered within the large-scale framework of cultural evolution, where social groups of non-kin may cooperate to ensure the survival and perpetuation of their culture over more individualistic goals. At a cultural systems level, a division of labor with each part holding a different potential utility or "cultural fitness" level may cause some individuals to sacrifice for others to attain better positions to defend or promote cultural

goals in a rationally justified way. According to this interpretation, devoted actors would be an adaptive product of cultural selection pressures in favor of cultural perpetuation (Kacou, 2012). However, such grand scope culture-level rationality is not incompatible with the fact that, at a psychological level, devoted actors rely on deontological, rule-bound or emotion-laden decision-making, pitting sacred values against individual interest. As Darwin intuitively points out, it is a high standard of morality which makes individuals ready to commit costly sacrifices for the sake of communal goals, wining over other competing tribes (Darwin, 1874).

1. 4. Neuroimaging studies relevant to SV research

1. 4. 1. fMRI studies on moral decision-making

Sacred values are defined as non-instrumental absolute moral values. As such, an appropriate starting point to inquire their neural underpinnings are neuroimaging studies on moral judgment, specifically those which make a distinction between neural processes underlying deontic and utilitarian moral decision-making. For instance, Greene, Sommerville, & Nystrom (2001) asked participants to judge the appropriateness of a series of moral dilemmas. Half of the scenarios appealed to the participant's personal involvement and, thus, elicited increased emotional processing (e.g. pushing a man to the railways to stop a train from killing 5 other people) and the other half were distinctively impersonal (e.g. pressing a button to deviate a train causing the death of 1 person in spite of 5). As the authors present, overcoming emotional reluctance to violate deontic rules (e.g. not to harm) for the greater good (e.g. saving 5 lives) characterizes utilitarian reasoning. According to their results, personal compared to impersonal moral dilemmas elicited higher activity in the medial frontal gyrus (mPFC), posterior cingulate cortex (PCC) and superior temporal sulcus/inferior parietal region (STS).

These three areas have been consistently related to social cognition in different studies. Lesions in medial aspects of the frontal gyrus have been observed to specifically relate to impaired socio-affective coding of one's own behavior, leading to socially maladaptive or "immoral" conduct, as in the famous case of Phineas Gage and the patient EVR (Bechara, Tranel, Damasio, & Damasio, 1996). The fact that IQ and abstract social knowledge was preserved in these patients leads Bechara and colleagues to conclude that the ventromedial

prefrontal cortex might sustain embodied simulations enquiring the affective valence of a given behavior in order to guide decision-making and attain adaptive social skills. Moreover, other studies point to a tendency towards aggressive violent behavior among individuals with damaged or reduced medial prefrontal cortex in a sample of Vietnam War veterans (Grafman et al., 1996), anti-social personality disorder patients (Raine, 2000) and a sample of criminals (Brower & Price, 2001). Further lesion studies associate vmPFC dysfunction with increased utilitarian reasoning in personal dilemmas (Young & Koenigs, 2007) and punitive reactions to unfair offers in the Ultimatum game (Koenigs & Tranel, 2008). Psychopathy has been also related to reduced prefrontal connectivity (Motzkin et al., 2012), an affective disorder that has been characterized by diminished empathic abilities (Blair, 2004), a preference for instrumental as opposed to reactive violence (Blair, 2001) and reduced limbic response to affectively loaded words (Kiehl et al., 2001). Conversely, in the general population, vmPFC circuits seem to respond more intensely to emotionally charged sensory stimuli (Rolls, 2000), during moral as opposed to non-moral social judgment (Moll, Oliveira-souza, Bramati, & Grafman, 2002), and in association with empathetic responses (Farrow et al., 2001). Thus, medial portions of the prefrontal cortex seem to incorporate affective significance into decision-making, a key function in representing others' mental states (Carrington, Bailey, Carrington, & Bailey, 2009) and, consequently, in deontic social moral judgment (Young, Cushman, Hauser, & Saxe, 2007). Additionally, Mitchell, Macrae, Banaji, & Hall (2006) establish a functional dissociation within the medial PFC, with ventral aspects more strongly related to self-referential processes, including others similar to the self, and dorsal portions encoding mental states of more distant others.

In turn, the superior temporal sulcus (STS) is thought to encode information about the intentionality of social agents (Frith & Frith, 1999), a system that presumably evolved to detect biological movement (Allison, Puce, & Mccarthy, 2000). Malfunction of the STS has been related to the Capgras delusion, where patients recognize their relatives physically but not affectively, accusing them of being impostors (Brothers and Ring, 1992). STS has been observed to respond to the movement of an agent's hands and eyes as well as to pictures of faces and bodies, which might be relevant cues to make assumptions about the intentions and beliefs of others (Borg, Hynes, Horn, & Grafton, 2006). Thus, the STS has been proposed to encode the affective and intentional aspects of social agents, a grounding element of theory of mind.

Finally, the posterior cingulate cortex (PCC) has been associated with moral versus non-moral judgment (Moll et al., 2002), adopting the first-person perspective (Vogeley et al., 2004) and processing of self-relevant emotional stimuli embedded within both imagined and autobiographical narratives (Summerfield, Hassabis, & Maguire, 2009). In sum, personal moral dilemmas seem to be recruiting brain areas involved in evaluating the affective significance (mPFC) of a social agent's intentionality (STS) possibly by referring to self-relevant narratives (PCC).

Conversely, in Greene's study (2001) impersonal dilemmas were more similar to non-moral dilemmas at a neural level, exhibiting higher activity in the dorsolateral prefrontal cortex (dlPFC) and the parietal lobe. The dlPFC is an executive function area typically related to working memory functions (Wager & Smith, 2003), logical thinking and cognitive inhibition (Koechlin, 2003; Ramnani & Owen, 2004). In interaction with affective processing, the dlPFC has been proposed to moderate detrimental emotions in social settings (Sanfey et al., 2003) and to overcome prepotent responses in favor of future abstract goals (Mcclure & Cohen, 2004). Thus, impersonal moral dilemmas that activate cost-benefit analysis strategies resulting in utilitarian outcomes involve brain structures involved in problem-solving and cognitive control (Greene & Haidt, 2002). This fact highlights the role of more executive rather than affective processes in impersonal in contrast to personal moral decision-making.

Taking personal moral dilemmas only, those which elicited longer reaction times ("difficult") were associated with anterior dlPFC, inferior parietal, ACC and PCC, when compared to those with shorter RTs ("easy") (Greene et al., 2001). Moreover, utilitarian decisions, i.e. those favoring utilitarian outcomes over deontic prepotent responses, were associated with longer RTs and neural activity in the same areas in addition to the temporal cortices in a follow-up analysis (Greene, Nystrom, Engell, Darley, & Cohen, 2004). Together with the brain areas commented above, the anterior cingulate cortex (ACC) has been identified as a cognitive conflict monitoring area, thus regulating goal-directed behavior (Ridderinkhof, 2004). According to the conflict monitoring hypothesis, ACC activity might be related to negative arousal states connected to changes in the autonomic nervous system (Critchley et al., 2003). Specifically, rostral parts of the ACC (rACC) have been associated with conflicting emotions (Etkin, Egner, & Kalisch, 2011) and cost-benefit ponderation (Rudebeck et al., 2008) necessary to overcome emotional

intuitions in favor of utilitarian outcomes. Strikingly, the rACC has been related to theory of mind and social information processing (Behrens, Hunt, Woolrich, & Rushworth, 2009), specifically encoding perceived social judgments of the self (Amodio & Frith, 2006) and feelings of guilt at unfulfilled social expectations (Zahn et al., 2009). Hence, difficult personal dilemmas where prepotent affective heuristics - usually derived from social expectations in ordinary settings - need to be disregarded for the greater utility (e.g. killing a crying baby to protect a group of people from being discovered by soldiers) might involve greater activity in executive function areas (dlPFC and inferior parietal cortex), as well as regions involved in social intentionality assessment (STS), self-referential narratives (PCC) and conflict monitoring (ACC).

Making use of the presented results, Greene et al. (2004), defend a model of moral decision-making where emotional automatic responses and cognitive control systems compete in favor of deontic and utilitarian outcomes, respectively. A distinctive aspect of Greene's studies is that personal dilemmas are designed to evoke deontic values specifically involving intentional physical harm to others. However, in a similar study, Kahane et al. (2012) used dilemmas involving a wider variety of deontic values comprising fairness, trust and not to kill, and found that deontologic compared to utilitarian judgments were associated with neural activity in the posterior cingulate cortex (PCC) and temporo-parietal junction (TPJ).

Apart from the PCC, commented above, the TPJ has been pointed out as a key region for theory of mind functions (Saxe & Kanwisher, 2003), including judgments of agency and awareness of the self (Decety & Lamm, 2007). However, it remains unclear whether the TPJ is involved in unspecific perceptual processes necessary for *mentalizing* (Mitchell et al., 2006) or, rather, it is a higher-level structure integrating ToM information aimed to predict and explain actions. The interplay between TPJ and PCC during deontic moral judgment has been proposed to sustain the evaluation of one's own actions with respect to self-relevant emotions, highlighting the presence of affective constrains guiding deontic reasoning (Kahane et al., 2012).

The same authors report higher activity in the insula, lateral orbitofrontal (OFC), ventrolateral prefrontal (vlPFC) and rostral anterior cingulate (rACC) during utilitarian in contrast to deontologic judgments (Kahane et al., 2012). The insula has been related with

negative emotions of disgust (Calder, Lawrence, & Young, 2001), *inequity aversion* (Hsu & Quartz, 2008), rejecting unfair offers in economic games (Sanfey et al., 2003) and care and compassion (Bernhardt & Singer, 2012). Meanwhile, the OFC has been involved in reward value coding of incoming stimuli (Y. Li, Vanni-mercier, & Isnard, 2016). Hence, it has been categorized as an executive function area selecting responses based on each option's reinforcement properties (Bechara, Damasio, & Damasio, 2000; Blair, 2004). Finally, the vlPFC or inferior frontal cortex is an executive function structure related to prepotent response inhibition both in the affective and motor domains, collecting information from emotional centers and enabling goal-oriented action (Sakagami & Pan, 2007). In sum, utilitarian judgment might comprise the assessment of global reward value (OFC), aversion towards unfair outcomes (insula), overcoming cognitive dissonances (ACC) and the suppression of prepotent affective responses (vlPFC).

In conclusion, moral judgment seems to involve both automatic emotional responses and cognitive evaluative processes. The first appear to be sustained by activity in self-referential nodes including the mPFC and the PCC and social cognition centers such as the TPJ and STS, generating affective intuitions that are generally aligned with social expectations and rely on inferences of others' mental states. The latter involve cognitive regulation structures such as the dlPFC, vlPFC and OFC, which inhibit maladaptive responses in order to maintain less immediate goals with higher reward value. In the face of difficult trade-offs, especially those compromising social expectations, conflict monitoring areas such as the rACC seem to come into play.

1. 4. 2. fMRI studies on sacred values

Only a few neuroimaging studies have specifically targeted sacred value processing: two of them under the moral theory framework (Berns et al., 2005; Duc et al., 2013) and one departing from the perspective of *vital loss* decisions (Li et al., 2011). Bern's and colleagues' experimental design (2012) was aimed to distinguish between utilitarian and deontic processing modes in response to sacred versus non-sacred values. First, participants were scanned while passively viewing of a series of statements encompassing mundane to increasingly sacred values (e.g. 'You are a dog person', 'You believe in God') and their complementary statements (e.g. 'You are a cat person'), after which they had to choose between the complementary options. Second, participants where asked whether

they would take any amount of hypothetical money to disavow their choice. This hypothetical phase was followed by an auction, where participants were required to ask for a price ranging from 1$ to 100$ to *sell* their chosen values in exchange for real money. This stage served to classify each statement as sacred or non-sacred for each participant, considering *sacred* those values which were non-negotiable and, thus, opted out of the auction. Finally, participants needed to sign a document including the values they had sold, thus testing the integrity of their hypothetical and factual willingness to sell each of their personal values. In the third stage, after a period of 6 to 14 months, participants were instructed to repeat their choices between complementary statements and report whether their selection had been a product of cost-benefit calculation (utilitarian choice), based on rights and wrongs (deontic choice) or neither.

As expected, the distribution of asked prices in the auction was bimodal, with values being sold either for very little (1$) or opted out, drawing a clear distinction between sacred and non-sacred values. SVs showed higher consistency (96.4%) than non-SVs in the follow-up survey and were more frequently categorized as "right and wrong" choices by respondents (73,2% versus 27,8% for non-SVs). In the neuroimaging analysis of the passive phase, the left temporoparietal junction (TPJ) was associated with statements justified as deontic choices in the follow-up survey, while the bilateral inferior parietal lobules (IPL) and the orbitofrontal cortex (OFC) were associated with utilitarian choices. The three resulting functional localizers were used to mask the contrast between SVs and non-SVs in the passive phase, with only the TPJ and IPLs showing significant activations for SVs and non-SVs, respectively. A whole-brain analysis of the same contrast revealed higher activity in the left vlPFC, dmPFC and right amygdala in response to items classified as SVs.

As the authors present, activity in the IPLs was significant both for non-SVs and cost-benefit choices, pointing to an involvement of these areas in utilitarian decisions involving mundane issues. On the other side, higher activity of the dmPFC and TPJ in association with SVs and "right or wrong" decisions is in line with moral judgment models that point to an intervention of ToM nodes associated with self and other's mental states in deontic reasoning, as presented above. In turn, amygdala activity in the SV condition might be related to higher arousal states or even perceived threat at openly visualizing personal values on the screen. Amygdala activity has been previously detected in the moral

judgment domain, especially in response to moral transgressions and when asked to judge the morality of an act (Moll et al., 2002). Finally, SVs were associated with ventrolateral aspects of the prefrontal cortices (vlPFC), thought to orchestrate behavioral responses combining information about reward, risks, probabilities and future goals (Hutcherson, Plassmann, Gross, & Rangel, 2012), a process which often results in suppression of prepotent responses. As proposed by Dixon & Christoff (2014), the vlPFC might be a key element in value-based decisions by representing associations between rules and motivational outcomes, thus constraining decision and action.

A further study by Duc and colleagues (2014) explores behavioral and neural responses to sacred value transgression making use of taboo trade-off scenarios with SVs pitted against instrumental values. Participants' responses were contrasted to scenarios involving tragic trade-offs between two conflicting SVs and routine trade-offs including two non-SVs. In order to evaluate such effects, two groups were built as a function of to their mean SV scores: a high-SV group, mostly consistent on theology students, and a low-SV group with mainly economy students. According to Duc and colleagues' results, taboo trade-offs were considered less acceptable than tragic or routine trade-offs, elicited negative emotions of moral indignation and were solved more easily, cutting off reaction times only in the high-SV group. This fact highlights the possible adaptive role of SVs as heuristic or rule of thumb in decision-making. On the other side, having to choose between two tragic outcomes was perceived as more stressful and difficult to solve, evoking greater feelings of uncertainty. A between-groups comparison revealed less acceptability, increased outrage and, at the same time, less difficulty and uncertainty in front of taboo trade-offs in the high SV group. The neuroimaging results established a link between SV violation (taboo trade-off) and bilateral amygdala and left anterior temporal lobe activation (aTL). In turn, activity in the right amygdala correlated with increased moral disgust and reduced uncertainty scorings in front of taboo trade-offs in high-SV individuals, who also exhibited higher neural activity in both right amygdala and aTL than low-SV individuals. According to these results, the aTL might be working as an integration node regulating emotional responses to taboo transgressions. The same area has been identified in social cognition tasks (Zahn et al., 2009) and morality studies (Kahane et al., 2012) and is thought to contain a supramodal integration center (Visser, Embleton, Jefferies, Parker, & Ralph, 2010). Moreover, a correlational analysis between different significant functional clusters in the taboo condition revealed a positive correlation between the left aTL and the right

amygdala, and between the left aTL and medial parietal areas (BA5). Additionally, the inter-cluster analysis revealed a significant negative correlation between the right dlPFC (BA8) and the left aTL, BA5 and, more prominently, the right amygdala. Importantly, such effects were only observed in the high SV group. The authors propose that the dlPFC could be working as a top-down regulating system enabling reappraisal of amygdala-driven emotional responses and redirecting bottom-up attentional switching mediated by medial parietal structures (Chun, Golomb, & Turk-Browne, 2011).

Finally, in a similar although not identical way, Li and colleagues (2011) presented participants three different scenarios involving personal vital loss (e.g. death of a child or being raped) and trivial loss, including *vital-vital*, *vital-trivial* and *trivial-trivial* loss choices. The concept of vital choice highly overlaps with that of sacred values, as it is defined as an inviolable and non-negotiable value. Behaviorally, and similarly to Duc et al. (2014), *vital-vital* loss choices were scored higher in difficulty and negative affect. Neurally, when compared to *trivial-trivial* choices, the *vital-vital* condition was associated with greater medial PFC, rACC, anterior medial TL, amygdala and PCC, even when controlling for decision difficulty measured by reaction times. Furthermore, activity in the amygdala and PCC correlated with negative affect and mPFC and rACC activations were associated with decision difficulty. As presented above, these five areas have been associated with moral judgment in different studies, highlighting once again the importance of emotional, self-referential and ToM processing during decision-making both in the vital and sacred domains. Moreover, Li et al. (2011) *vital-trivial* loss choices were associated with positive affect responses and higher activity in the OFC and ventral striatum. Such results were interpreted as feelings of relief to the fact that choosing trivial loss exempted them from having to incur in vital loss. Thus, the *vital-trivial* condition in this study is not directly comparable with Duc's taboo trade-off condition, which was rather associated with negative moral outrage responses. The antagonical emotional responses to the mixed condition in both studies could be due to the different decisional framings: whereas participants in Li et al. (2011) were given the chance to elude a vital loss by choosing a trivial loss, Duc et al. offered a trivial gain in compensation for a sacred loss, thus critically devaluating sacred concessions.

As reviewed in this section, some neuroimaging studies to date have identified neural activity in brain regions such as the vlPFC, the TPJ and dmPFC in association with SVs

and medial PFC, anterior temporal lobe and amygdala activations in decisional settings involving SVs. However, much work is left to do with respect to the motivational aspects of SVs, i.e. how moral community members are motivated to defend SV integrity, and what is the role of the social context in brewing and maintaining such values over time. Whilst the first question remains mostly opaque to date, a solid body of neuroimaging research has been addressing the impact of social input on personal values and attitudes, which can be helpful in understanding how a moral community influences its members' commitment to group-relevant SVs. The following section reviews social conformity literature relevant to the present study.

1. 5. Neuroimaging studies on social conformity

Acquiring cultural knowledge by imitating others is an adaptive learning strategy favored by natural selection within subpopulations with a certain degree of migration (Boyd and Richerson, 1985). As such, deviating from popular behaviors involves a net cost, while conforming to the norm entails net benefits in terms of biological fitness at a population level. In a culturally complex world, social conformity fulfills not only normative goals by providing approval within a social community, but also informational goals. An historical example of socially compliant behavior motivated by social approval seeking is found in Asch's experiment (1955). In it, participants changed their judgment on a simple perceptual task after learning that their peers disagreed with them. Conversely, *blind* informational reliance on the group is proven by the persistence of obviously incorrect socially influenced choices even in the absence of peers (Deutsch, 1955). Interestingly, Morgan & Laland, (2012) found that only when the number of social demonstrators is high enough and an individual's certainty about her own abilities is low, social conformity responses begin to increase. Moreover, developmental studies on social conformity show that, when it comes to adopting social information, children rely more heavily on individuals that resemble members of their own cultural group (Corriveau, Fusaro, & Harris, 2009). In turn, selectiveness in trusting a social source increases with age, filtering out testimonies that contradict one's own experience (Clément, 2010). Later, during adolescence, presence of buddies amplifies an already enhanced responsiveness to reward (Chein, Albert, Brien, Uckert, & Steinberg, 2012), suggesting higher sensitivity to social influence during this period.

In fact, susceptibility to social inputs has been related to general biological sensitivity to context (Ellis & Boyce, 2008), as it seems to respond to reinforcement brain systems encompassing both reward seeking and pain avoidance behaviors. Four key brain structures associated with these processes have been related to the social conformity bias. Thus, whereas the ventromedial prefrontal cortex (vmPFC) and the ventral striatum (VS) mediate reward-related responses (Haber & Knutson, 2009), the dorsal anterior cingulate cortex (dACC) and the anterior insula (AI) are involved in detection of pain both in the physical and the social domain, as in the case of social exclusion (Eisenberger, 2012). Concretely, the vmPFC has been observed to exhibit higher activity in response to social pressure (Izuma, Saito, & Sadato, 2008) and is thought to integrate socially relevant information and personal internal drive to accomplish goal-directed behavior (Cunningham & Zelazo, 2007). In turn, more dorsal aspects of the mPFC have been related to the assessment of social relevance (Mason, Dyer, & Norton, 2009). Moreover, neural activity in the ventral striatum has been reported to increase at conforming social opinions (Campbell-Meiklejohn, Bach, Roepstorff, Dolan, & Frith, 2010) and decrease at non-conforming opinions (Klucharev, Hytönen, Rijpkema, Smidts, & Fernández, 2009). Such observations point to the social modulation of VS activity, ultimately impacting reward value perception. Conversely, the dACC exhibited increased activation at non-conforming opinions (Klucharev et al., 2009), and, together with the AI, its activity predicted change of judgment to conform to the social norm (Berns et al., 2005). Thus, both structures have been hypothesized to encode the costs of social exclusion (Juvonen and Gross, 2005).

At a molecular level, three main neurotransmitter systems regulate neural activity in social reinforcement centers: dopaminergic activity regulates both the VS and the dACC (Vollenweider, Vontobel, Hell, & Leenders, 1998), serotonine mediates reward responses by affecting the VS (Kranz, Kasper, & Lanzenberger, 2010) and opioids modulate pain responses through the VS, ACC and insula (Svingos & Garzo, 2001). Consistently, psychological genetic studies have revealed a series of associations between genetic variants and responses to social reinforcement. On one hand, variants linked with higher dopaminergic function such as the *Met* allele of the *COMT* enzyme has been associated both with higher vmPFC reward-related activity (Dreher, Kohn, Kolachana, Weinberger, & Faith, 2009) and processing of negative affect (Mier, Kirsch, & Meyer-lindenberg, 2010). Similarly, the *MAOA*-uVNTR allele is linked to dopaminergic and serotoninergic

dysregulation (Shih et al., 1999) and has been associated with heightened ACC processing (Buckholtz et al., 2008) and increased susceptibility to social exclusion (Eisenberger, Way, Taylor, Welch, & Lieberman, 2007). Finally, the G allele of the opioid receptor gene *OPRM1* has been related to increased ACC and insula activity and heightened sensitivity to physical pain (Tan, Lim, Teo, Lim, & Sia, 2009) and concern about social exclusion (Way, Taylor, & Eisenberger, 2009). Additionally, in a pharmacological study, increased dopaminergic function through administration of methylphenidate (Volkow et al., 2001) increased conformity with the group's opinion (Campbell-Meiklejohn et al., 2012) after moderate but not large social dissonance. According to these observations, catecolaminergic and opioid transmission play a decisive role in mediating individual responses to both physical and social reinforcement cues, ultimately impacting an individual's disposition to conform to social norms (see Falk, Way, & Jasinska, 2012, for a more extended review).

fMRI studies using subjective attractiveness ratings ratify the role of social input in reinforcement learning, which not only affects subsequent public ratings but also subjective value attribution observable at a neural level (Zaki, Schirmer, & Mitchell, 2011). In a study run by Klucharev et al. (2009) female participants were instructed to rate the attractiveness of a series of female faces while being scanned. After each stimulus rating, a fictive average European rating was presented in a red frame, which was 2 or 3 points above or below the participant's rating in two thirds of the items and coincided with it the other third. At the end of the scanning session, participants were asked to rate the same faces in a second round, this time without group ratings. According to their results, opinion alignment with the social norm was significant. Neurally, conflict with normative group opinion induced higher activity in the rostral cingulate zone (RCZ), precuneus, dlPFC (BA9) and insula, whereas activity in the nucleus accumbens (NAc), dlPFC (BA6) and PCC was reduced. A conjunction analysis revealed that the RCZ hyperactivity and NAc hypoactivity associated with social conflict also predicted subsequent change of judgment to conform the social norm. Additionally, participants who scored high on conformity exhibited a stronger deactivation in the NAc when facing social conflict, in contrast to low-conformity scorers. In turn, the former required weaker conformity-related activity than the latter to align to the social norm in the subsequent rating. As the authors present, the RCZ seems to play an active role in reinforcement learning by creating a prediction error signal representing the difference between predicted and obtained

outcomes, which is then used to adjust future decision-making (Cohen & Ranganath, 2007). Meanwhile, the NAc seems to respond more promptly to unpredictable reward (Doherty, 2004) and reward cues, thus improving reward prediction abilities during reinforcement learning (Knutson & Wimmer, 2007). In a methodologically similar experiment, Zaki et al. (2011) scanned participants during the second attractiveness rating, after receiving normative peer opinion for each face, in order to capture brain activity differences in reward-related circuits. Their work shows not only that participants conformed to the group after receiving normative opinion, but also that such changes corresponded to organic shifts in OFC and NAc responsiveness. Such pattern of activity points to subjective changes in reward value coding in accordance to the social norm, which spanned across the public and the personal domains. This finding endorses the notion of a sociobiological adaptation aimed to reduce group deviations and motivate consensus, thus aligning decision-making and coordinating behavior at a group level.

Finally, Pincus, LaViers, Prietula, & Berns (2014) investigated the role of social influence on personal attitudes specifically concerning sacred values. Making use of an experimental design similar to Berns et al. (2012) including a passive phase, a hypothetical phase and an auction, where participants proposed ask prices to sell their previously chosen values, as described in the previous section. However, in this study, social group opinions were presented to participants before being asked about their hypothetical willingness to disavow their personal values in exchange for money. Normative opinions were presented in form of the percentage of in-groups agreeing with the values or in form of written statements (e.g. “One third of the participants agreed with you”). Intervention format did not affect results. In contrast to Klucharev (2009) and Zaki (2011), ventral striatum did not show socially modulated neural activity changes, which according to Pincus (2014) might be due to the lack of hedonistic value of the personal values at stake. In turn, left vlPFC activity positively correlated with those values which were opted out from the auction (sacred values), thus replicating Berns et al. (2012) findings. Interestingly, vlPFC activity was weaker in participants who were more sensitive to social influence in the auction phase, rising prices or opting out for values with increased social popularity. Authors portray the left vlPFC as a rule retrieval node, whose activity correlates with increasing norm rigidity and resistance to belief revision, ultimately precluding change in response to social contingencies. Consistently, a study by Sharot et al., (2012) found that disruption of specifically left vlPFC by means of transcranial magnetic stimulation promoted belief

updating. Nonetheless, Pincus et al. (2014) provided real group ratings, which were not normally distributed, thus skewing the data towards non-conflicting normative opinon. In addition, although social popularity of two competing complementary values was assessed, group's average *degree* of endorsement was not acquainted. Such approach would have required the use of a quantitative metrics, as in Klucharev (2009) and Zaki (2011), designed to measure degree of acceptance of a positive condition (e.g. facial attractiveness). Hence, while Pincus et al. (2014) measured the effects of value *popularity* on value sacralization, how the *instensity* of personal SV endorsement is affected by confederates' ratings on the same measure remains unclear.

2. Outline, objectives and hypotheses

2. 1. Outline

As reviewed so far, sacred values (SVs) are absolute moral imperatives that play a striking role in defining an individual's social identity. Importantly, SVs are key elements in sustaining intractable inter-group conflict over time, a phenomenon that has been observed, for example, in contemporary conflicts (Dehghani et al., 2010; Ginges et al., 2007; Sachdeva & Medin, 2009). Field studies on SVs have provided some insight into how people behave and decide with respect to their SVs. On one side, people exhibit moral outrage reactions in the face of transgression or instrumentalization of their values. On the other side, in situations of threat, confederates strengthen their commitment to their in-groups, with whom they share specific SVs. Group commitment is often nourished through the inclusion of such values both in community rituals and into narratives involving national or religious identity. Such mix of factors promotes an environment where violent attacks and costly sacrifices in defense of group-related SVs are more likely, often escalating in a chaotic manner. Observations of the like suggest that, especially in situations of deep-seated conflict, SVs must receive a special treatment in order to enable negotiations. For instance, mutual acceptance of peace deals is more likely when, rather than one-sided economic offers, bilateral concessions in the sacred domain are made. Simultaneously, in a more theoretical ground, the existence of SVs challenges classical utilitarian models of decision-making in favor of the devoted actor model, which portrays decision-makers who rather rely on deontic rules responding to social identity constraints sculpted by social expectations.

In the field of neuroimaging, however, very few studies have specifically targeted SVs and the neural effect of environmental factors on SV adherence, such as social influences. Similarly, no studies to date have investigated the neural mechanisms underlying personal motivation to engage in violent action such as fighting and dying in defense of group-specific SVs. Therefore, the present research aims to contribute to the very recent but growing corpus of evidence characterizing neural activity patterns involved in SV processing making use of an exceptional real-world sample of Pakistani males with radicalized political views. Concretely, this study aims to elucidate 1) how SVs activate

motivational brain centers in order to induce behavioral outputs aimed to secure and promote group-related SVs (Experiment 1), and 2) how SV-related neural circuits may respond to specific social inputs conveying different degrees of support for SVs (Experiment 2). In this pursuit, and taking advantage of the more extensive literature exploring moral cognition and social conformity, the present study makes use of a false consensus feedback paradigm, similar to the ones used by Klucharev (2009) and Zaki (2011), designed to explore the effect of social normative opinion on personal attitudes, in this case, on willingness to fight and die in defense of SVs. This approach allows capturing both the neural mechanisms underlying motivation to protect SVs at the cost of one's own life (Experiment 1) and the effect of the in-group's normative opinion on such circuitry (Experiment 2). In order to study these processes in the context of real inter-group conflict, the participants' sample consists of 30 Muslim Pakistani males recruited in the area of Barcelona supporting the Kashmiri cause and scoring high in a construct representing political radicalization (see Methods, section 4.1.)

2. 2. Objectives

2. 2. 1. General objective

The aim of the present study is twofold: i) pinpointing the neural mechanisms of willingness to fight and die for sacred values (Experiment 1) in a radicalized sample of Muslim Pakistani males and ii) assess the behavioral and neural effects of social influence on the same measure (Experiment 2) in the same sample.

2. 2. 2. Main objectives

The main objectives of this research are:

1. To identify neural mechanisms of willingness to fight and die in defense of sacred values compared to non-sacred values in a group sample endorsing political violence (Experiment 1).
2. To detect brain activity associated with the reception of social normative opinions on willingness to fight and die for values that are sacred to the participant (Experiment 2).
3. To evaluate whether social normative opinion has an impact on participant's willingness to fight and die ratings (Experiment 2).

2. 2. 3. Secondary objectives

1. To detect behavioral discrepancies in willingness to fight and die for SVs in contrast to non-SVs.
2. To assess the effect of conflicting social normative opinion on change of judgment in willingness to fight and die for SVs compared to non-SVs.
3. To evaluate whether neural activity associated with social conflict predicts changes in judgment in willingness to fight and die for SVs and non-SVs.
4. To examine whether self-reported moral outrage associated with conflicting social normative opinion predicts activity in social conflict brain areas.
5. To identify variables related to in-group and out-group perceptions that can help predict neural activity associated with willingness to fight and die SVs.

2. 2. Hypotheses

Based on previous behavioral and neuroimaging studies encompassing moral judgment, sacred value processing and social conformity we hypothesize that:

H1a. SVs will elicit higher and faster willingness to fight and die ratings than non-SVs.
H1b. Brain areas previously associated with SVs and deontic decision-making such as the left vlPFC (Berns et al 2012) and the vmPFC (Greene, 2015) will be more active in the SV condition, whereas areas related to cost-benefit calculation such as the dlPFC and bilateral inferior parietal cortex (Greene et a., 2001) are expected to be more active in the non-SV condition.

H2a. Conflicting in contrast to non-conflicting normative opinion is expected to induce higher activity in prediction error regions such as the rostral cingulate zone (Klucharev et al., 2009) and areas associated with disgust and moral outrage such as the insular-opercular complex (Hayes & Northoff, 2011).
H2b. Neural activity differences associated with conflicting normative opinion presented in H2a will be enhanced in the sacred value condition compared to the non-sacred value condition.

H3a. SVs are hypothesized to be more resistant to social influence, thus change of judgment after receiving normative opinion is expected to be lower in SVs compared to non-SVs.
H3b. Change of judgment is expected to predict increased neural activity in areas responsive to conflicting normative opinion (H2a).

H4a. SVs are expected to elicit higher moral outrage responses in front of conflicting normative opinion than non-SVs.
H4b. Moral outrage ratings are expected to predict neural activity in brain regions processing moral emotions of disgust such as the insular-opercular cortex.

3. Methods

3. 1. Participants

Thirty Pakistani Muslim males between 18 and 39 years old (mean age = 29.83 yrs, sd = 5.51) were recruited for the study. Selection of participants was conducted by means of an interview on the field in different neighborhoods around Barcelona with a high density of immigration, e.g. 49% of foreign population in El Raval, with Pakistanis being the most numerous population (B-raval Iniciatives de Solidaritat i Promoció, 2015). The field interviews were elaborated by a team of Urdu speaking experimented field researchers who had been building relationships with cultural and religious leaders as well as average population in the Pakistani community in Barcelona for two years (2014 - 2016). During the interview, candidates' opinions on several topics related to Islam were evaluated (see full survey document in Annex). From a sample of 142 interviewed people, a group of 30 participants holding at least 6 sacred values and 6 non-sacred values and scoring high in a construct representing political radicalization were included in the neuroimaging study (criteria for sacred value selection and political radicalization follows).

Sacred values were chosen from a list of 25 candidate values that was shortlisted by means of previous field testing with Pakistani males by the same team of field researchers (e.g. "India should have no rights on Kashmir", "US military should be expelled from Muslim lands"). A given candidate value was considered sacred for a particular participant only if the following conditions were fulfilled: 1) the participant rejected any financial concession in exchange to give up his position making use of case-based scenarios, 2) the participant was not willing to accept the contrary position even if there was a democratic consensus over that issue. Conversely, a value was considered non-sacred if it failed to meet any of both of the above conditions. The full list of the candidate values and case-based financial concession scenarios can be found in the survey (see Annex).

Political radicalization was scored by means of a construct encompassing: 1) support for the militant group Lashkar-e-Taiba (see Box 3) involved in the conflict over Kashmir (see Box 2) with a score of at least 4 in a 5-point Likert scale, 2) agreement with the statement "Armed jihad should be waged against those who seek to harm Muslims", 3) willingness

to incur in at least 1 costly sacrifice involving violence in defense of any of the chosen SVs, including "violently protest to encourage my country to go to war", "support financially or in spirit a jihadi group", "join a jihadi group" and "fight and die on my own".

Exclusion criteria included comorbidity with psychiatric diseases or personality disorders, assessed by the MINI International Neuropsychiatric Interview (Sheehan et al., 1992). Only participants with good Urdu reading and comprehension skills were included. The study was approved by the Institutional Review Board (IRB) Ethics Committee and informed consent was obtained from all participants before taking part in the study. Data protection was a top priority and complete anonymity was guaranteed.

From the 30 selected participants, half of them were married, twenty-two reported having "well below the average income" in Spain and all but five had attained at least a high school diploma. Twenty-eight of them were Sunni, one Deobandi and one Ahle Hadees. Private prayer and mosque attendance was once or twice per week at average. Twenty-two of them stated that religion was "the most important thing in their lives". All of them agreed that "India should have no right to the ownership of Kashimir".

Box 2. The Kashmir conflict

The Kashmir conflict started in 1947 with India's independence and its partition from Pakistan, and has persisted over 60 years without arriving to any satisfactory agreement. Although it began as a territorial dispute it has evolved into a clash between the Hindu and Muslim religious identities (Rai, 2004), specially salient for the Muslim community (Sachdeva & Medin, 2009).

Despite the Standstill Agreement signed by the Kashmir Hindu Maharaja in 1947, the majoritarian Muslim population revolved and invaded Kashmir causing mass killings of all ethnicities and hundreds of thousands of refugees. The Maharaja put Kashmir under Indian control and the UN Security Council imposed a ceasefire, forcing Pakistan to withdraw its troops, and promised a plebiscite for Kashmiris to democratically vote for

their independence. Neither Pakistan obeyed nor the plebiscite took place, creating a climate of escalating violence leading to the Sino-Indian war in 1962 - which terminated with the Chinese occupation of the eastern territory - and the Indo-Pakistani war in 1971 – that resulted in Pakistan giving up control over the current Bangladesh and the definition of the Line of Control between India and Pakistan (see Figure 2). Indian military occupation lead to the formation of opposing Muslim militant groups, culminating in the Mujahedeen insurgency in 1989, which has caused 70000 deaths and 8000 missings to date. Many of the insurgent groups identify with and receive support from Jihadist movements including Lashkar-e-Taiba, one of the three main Muslim militant groups.

Figure 2. Map of the Kashmiri territory and the areas of controlled by Pakistan, China and India. Adapted from the Toronto Star Graphic.

The most recent ceasefire resolution between Pakistan and India was signed in 2003, although there have been several violent attacks from both sides ever since, such as the 2008 Mumbai attacks perpetrated by Pakistani militants. Currently, 43% of the territory is controlled by India, 37% by Pakistan and 20% by China. The general Pakistani view defends that Kashmir should be independent or annexed to Pakistan, according to its Muslim majority and responding to the Kashmiri uprising. In turn, Pakistan accuses India of having disavowed the UN resolution of holding a plebiscite and makes the Indian military responsible of thousands of deaths and rapes during the nineties.

Source: http://www.cfr.org (Council on foreign relations) and www.insightonconflict.org

Box 3. *Lashkar-e-Taiba*

Lashkar-e-Taiba (LeT, or the "army of the pure") is one of the largest militant Islamist groups in Pakistan functioning since 1993 as the military wing of Markaz-ad-Dawa-wal-Irshad, an Islamist organization dedicated to recruit fighters in support of the Taliban since 1989 (see emblem in Figure 3). They have been attributed several violent attacks including the 2008 Mumbai bombings, which caused 200 deaths and 300 injured civilians, and the Indian parliament attacks in 2001, accusations which the group denies. It has been estimated that the organization counts on several thousands of militants in the Kashmir, most of them non-Kashmiris from Pakistani and Afghan origin, according to the US state department. As reported by the South Asia Terrorism Portal (SATP), the group controls 135 secondary-schools and 16 Islamic institutions, and its arsenal includes assault rifles, machine guns, explosives and grenades. They disseminate their opinions in an weekly Urdu publication, *Gazwa*, and a monthly English-language magazine, *Voice of Islam*, among others. Funds are seemingly collected from Pakistani and Kashmiri businessmen, expatriates living abroad, Islamic non-governmental organizations and, allegedly, Saudi Arabia and the Pakistani intelligence.

The group's ideology does not only include the liberation of Kashmir from Indian control, but also the unification of all Muslim majority regions beyond Pakistan, and points at the US, Israel and India as "existential enemies of the Islam", as reported by the SATP. Meanwhile, the US (2002) followed by the UN (2005) officially declared the LeT a terrorist group, which lead to the prohibition of the organization in Pakistan in 2002. Ever since, the organization is thought to have adopted the name of *Jamaat-ud-Dawa* (JUD, the society for preaching), founded on behalf of the victims of the Kashmir earthquake in 2005. The US State Department reported that schools in control of JUD work as recruitment centers for Islamic fighters and the organization has cooperated with al-Qaeda at least in some cases. Experts point out that LeT continues to be active in Pakistan and has emerged as a global terrorist group operating in Afghanistan, Chechnya and Iraq.

Figure 3. Lashkar-e-Taiba emblem adapted from www.indiandefensenews.in

Source: http://www.cfr.org (Council on foreign relations) and www.insightonconflict.org

3. 2. Neuroimaging paradigm

A rapid event related fMRI paradigm with randomized trials and jittered inter-trial intervals (ITI) was designed using Matlab with Psychtoolbox extensions (Kleiner et al, 2007). The items were presented in Urdu language, translated by an Urdu linguist and revised by a group of Urdu speaking collaborators. From the list of sacred and non-sacred values elaborated for each participant in the field survey, 6 randomly selected non-SVs and the 6 top SVs ranked in order of importance by the participant were included in the paradigm. Thus, each participant completed an individualized version of the paradigm including their own collection of sacred and non-sacred values. Once arrived to the fMRI facilities and instructed for the experimental task, participants practiced a training version of the paradigm on a laptop with the same structure but different values with the guidance of two Urdu-speaking assistants for about 10 to 15 minutes. The fMRI session included two experimental paradigms: *Rating1* and *Feedback*, each of them was split in 2 runs. The whole fMRI session took about forty minutes to complete.

The *Rating1* paradigm (Experiment 1) included a total of 80 trials, each of them including one of the participant's selected values. In order to cover the 80 trials, values were presented in different grammatical forms randomly selected from a previously developed collection of rephrasings, including seven formulations per value controlled for sentence length and semantic content. The list of rephrasings was elaborated by a professional Urdu linguist and revised by Urdu-speaking collaborators. Participants were instructed to convey their willingness to fight and die for each of the presented values making use of a

Likert scale ranging from 1 (not at all willing to fight and die) to 7 (extremely willing to fight and die), as illustrated in Figure 4.

The *Feedback* paradigm (Experiment 2) included 120 items, encompassing the 80 items from *Rating1* and 40 randomly chosen rephrasings of the same values (20 sacred and 20 non-sacred) (see Figure 4). In this block, participants witnessed their own previous rating framed in green followed by the normative opinion of the Pakistani community in Barcelona in red. In order to keep participants attentive, they were instructed to press a button to proceed to the community feedback screen. Community ratings were manipulated to be 2-points less willing to fight and die (*peers-lower* condition), non-conflicting with the participant (*peers-agree* condition) and 2-points more willing to fight and die (*peers-higher* condition) in the same proportion of trials in order to obtain enough statistical power, i.e. approximately 20 trials for each of the 3 feedback conditions x 2 sacredness conditions (SVs/non-SVs). Each value received consistent feedback in all of its rephrased forms. Extreme ratings could only be included either in the *peers-agree* condition or in one of the conflicting conditions (*peers-higher* for ratings 1 and 2 and *peers-lower* for ratings 6 and 7). Given that SVs generally received ratings in the higher extreme, a bias was generated with SVs hardly ever receiving higher peers' ratings.

Immediately after the scanning paradigms and without previous notice, participants were asked to complete a third task, Rating 2, on a laptop and sitting in a separate room. In it, participants had to convey their willingness to fight and die for each of the 12 values included in both scanning paradigms over again in order to detect possible changes of judgment (rephrasings for the values were excluded at this point).

In the final stage, participants completed an *emotions* questionnaire. In it, they were first reminded of their own ratings in *Rating1* along with their peers' ratings in the *Feedback* section, and then were asked about the emotions they experienced at witnessing the Pakistani community's normative opinion. The emotions included disgust, anger, contempt, shame, joy, compassion and pride in a 7-point Likert scale. Disgust, anger and contempt scores were collapsed into one single construct labelled "moral outrage".

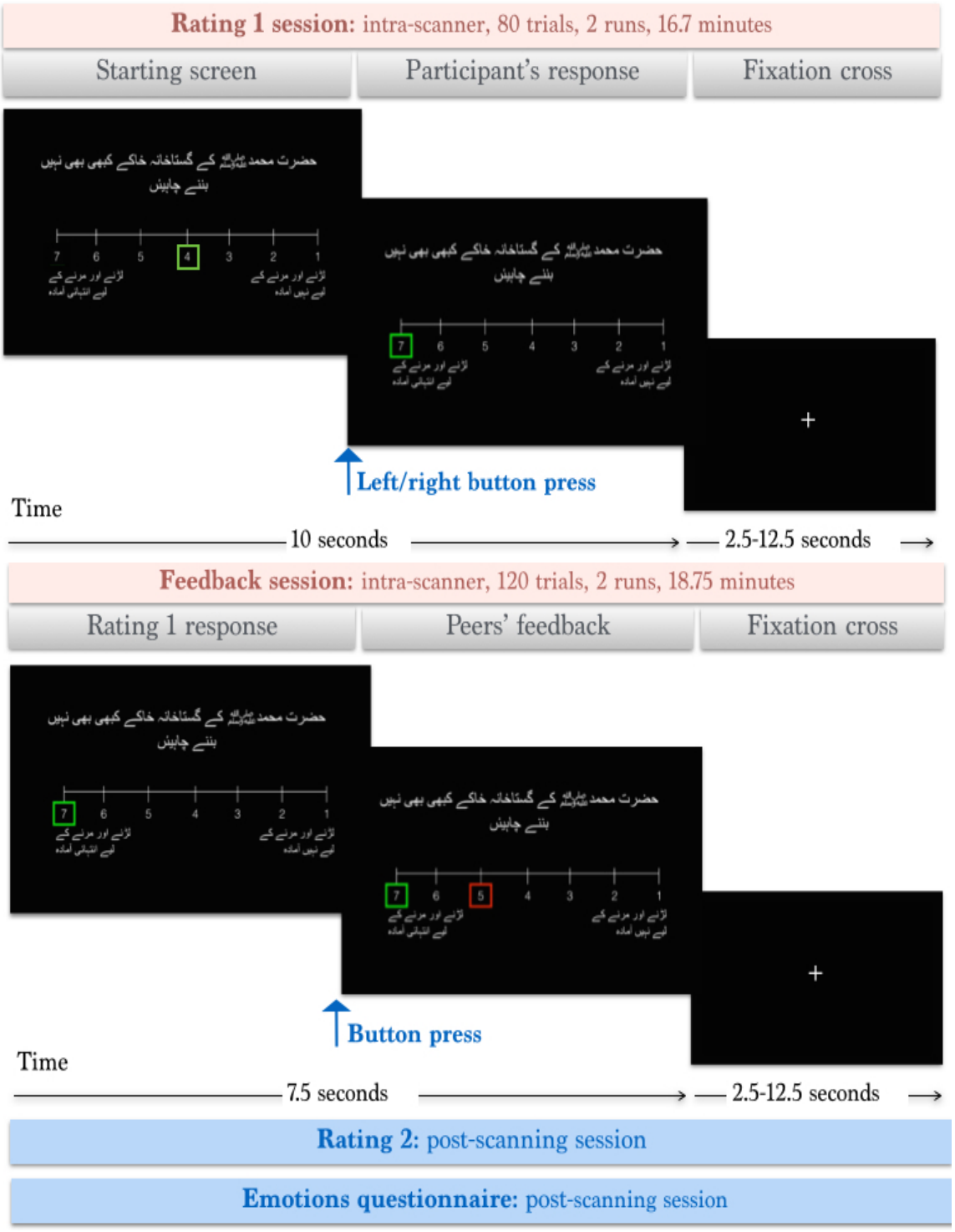


Figure 4. Scheme of the overall testing timeline including trial structure of Rating 1 and Feedback paradigms. Trials are exemplified with the screenshot of one of the presented items in the Urdu version ("Prophet Mohammed must never be caricatured") with a Likert Scale ("Not at all willing to fight and die" to "Extremely willing to fight and die"). Rating 1 trials started with the cursor centered in 4 (framed in green), which participants moved along the Likert scale to convey their rating. Feedback trials started with the presentation of previous participants' ratings on the same item (framed in green). Community feedback (framed in red) was presented following button press. After completing the Rating and the Feedback paradigm, a behavioral testing session including Rating 2 and the emotions questionnaire was completed outside of the fMRI scanner.

3. 3. Analysis of behavioral data

In the analysis of *Rating1*, two repeated measures ANOVA were conducted with trial type (sacred vs. non-sacred values) as a within-subjects factor to predict both willingness to fight and die ratings and reaction times.

Data from the *Feedback* session was combined with *Rating 1* and *Rating2* scores in order to analyze change of judgment as a function of feedback condition. For this purpose, a third repeated measures ANOVA was performed assessing the effects of i) trial type (Sacred vs. non-Sacred values), ii) feedback condition (*peers-higher, peers-lower* and *peers-agree*) and iii) time (Rating 1 vs. Rating 2) on self-reported willingness to fight and die for a given value.

Finally, a series of repeated measures ANOVAs were conducted in order to elucidate the effects of i) trial type (Sacred vs. non-Sacred values) and ii) feedback condition (*peers-higher, peers-lower* and *peers-agree*) on self-reported emotions towards the in-group (moral outrage, shame, joy, compassion, pride) after receiving community ratings. Correlational analysis between willingness to fight and die ratings, change in judgment and emotion scores were performed.

On a separate note, the seventh item in the costly sacrifices questionnaire “Carry out militant actions to fight and die on my own, even if I had no group or support network” (see annex, CS7) was explored in relationship to all other measures taken in the survey, as it was one of the most relevant questions in terms of radicalization.

3. 4. Neuroimaging data acquisition and analysis

Images were acquired in a Siemens 3T scanner. T1-weighted images were obtained using a FSPGR sequence (TR: 11.6 ms, TE: 4.8ms, FA: 12, matrix size: 280 x 280, 150 slices, slice thickness: 1.00mm). An EPI-T2* sequence allowed obtaining the functional volumes, each comprising forty 3.4 mm thick slices (TR 2500 ms, TE: 27 ms, FA: 90, matrix size: 64 x 64, 40 slices, slice thickness: 3.4 mm).

For the functional MRI data analysis, the software package SPM12 (Welcome Department of Imaging Neuroscience, London, United Kingdom) and the Analysis of Functional Neuroimaging software (AFNI, Cox 1996) were employed. Functional images were despiked (AFNI), corrected for motion-related artifacts (SPM), normalized to MNI standard space (SPM), smoothed with a 12mm full-width-at-half-maximum Gaussian kernel (SPM) and detrended (AFNI).

3. 4. 1. Univariate analysis

The univariate analysis was conducted using the software package SPM12 (Welcome Department of Imaging Neuroscience, London, United Kingdom). In the first-level analysis for the *Rating1* paradigm (Experiment 1), a GLM matrix design was modelled including 2 Sacredness regressors (Sacred/non-Sacred values), a fixation cross regressor, and an extra "inattention" regressor modelling both non-responded trials and odd responses. Odd responses were defined as rating scores that were more than 2-points below or above the median score of all the rephrasings derived from the same value. Rephrasing scores for the same value were broadly consistent, with odd responses adding up to an average of 1.46 per participant and a maximum of 8 out of the 80 trials. Six movement regressors were included. Contrasts of interest included neural activity during sacred value ratings compared to non-sacred value ratings, and activity of both conditions against baseline activity (fixation cross).

The GLM matrix for the *Feedback* paradigm (Experiment 2) encompassed 6 regressors (3 feedback types x 2 sacredness conditions), a regressor for the fixation cross, and an extra "inattention" regressor modelling non-responded trials and odd feedback trials. Odd feedback trials took place when the social normative rating was over two points above or below the participant's previous rating. This was only the case if participants' responses in *Rating 1* had been odd (mean = 1.46 per participant). In that case, instead of calculating the feedback based on the participants' previous rating in a given rephrasing, it was calculated based on the median rating of the value, thus ignoring *Rating 1* odd responses. This strategy helped maintain a small variance across the social feedback for all rephrasings stemming from the same value in favor of a higher consistency, and therefore credibility, of the feedback. Finally, six movement regressors were included in the GLM. In order to capture neural activity associated with the peers-lower feedback condition for

sacred values (-2 points in the 7-point willingness to fight and die Likert scale), it was contrasted to i) peers-agree feedback for sacred values (0 points difference with respect to participant scores) and ii) peers-lower feedback for non-sacred values. Interaction effects between feedback type and value sacredness were tested.

In addition, in order to capture individual differences in neural activity across the whole group, four multiple regressions were performed. Two regressions included out-group perception measures obtained in the field survey, i.e. US spiritual and physical formidability scores, as predictors of neural activity associated with willingness to fight and die for SVs in the *Rating 1* paradigm. The spiritual formidability measure assesses perceived spiritual commitment of the members of a given group or country (in this case, the US) represented as the size and strength of a male body, whereas the physical formidability measure represents the potential of a group as a rival in terms of material resources (see Annex, questions 33 to 36). The other two regressions included 1) degree of change in judgment after peers-lower feedback for SVs and 2) self-reported moral outrage at *peers-lower* normative opinion for SVs as predictor of neural activity associated with peers-lower feedback for SVs in the *Feedback* paradigm. In each contrast, a whole-brain exploration was conducted using a threshold of $p<0.05$ corrected for multiple comparisons by means of the Family-wise error rate.

3. 4. 2. Psychophysiological Interaction Analysis

The psychophysiological interaction analysis was performed in SPM12 (Welcome Department of Imaging Neuroscience, London, United Kingdom). It allows finding synchronized activity between two brain regions triggered by a specific psychological context. Such analysis helps build simple effective connectivity models based on regressions where activity in one brain area contributes to the activity of a different area in a given experimental condition. Unlike a typical functional connectivity analysis, PPI allows to distinguish the directionality of the influence of one area with respect to the other.

Based on the presented hypotheses, we were interested in evaluating whether brain regions typically associated with utilitarian decision-making such as dorsolateral prefrontal cortex (dlPFC), which we expected to be more active during the non-sacred value condition,

worked as inhibitory centers on neural nodes associated with the sacred domain. Conversely, we tested whether brain regions associated with sacred value assessment were inhibiting areas related to deliberative thinking (dlPFC).

In order to carry out this analysis, regions of interest such as the dlPFC and the vmPFC, expected to be differentially activated during the sacred (SV) compared to the non-sacred value (non-SV) condition, were selected as seed regions. In the first place, a mask encompassing the region of interest (e.g. dlPFC, vmPFC) was built and projected onto the participants' brains during the contrast non-SV vs. SV using the *Rating1* univariate analysis GLMs (see section 4.4.1). From the resulting matrix of ROI BOLD activity, the voxel of maximum activity over the threshold $p < .05$ was identified for each individual. A subsequent volume of interest (VOI) shaped as a sphere with a radius of 6mm and centered in the identified voxel of maximum activity was built for each participant. This way, participants counted on an individualized seed region with maximum activity in the region of interest. In the next step, the mean BOLD activity of the VOI in the contrast non-SV vs. SV was extracted for each time point and stored as a vector. This vector was then added as a regressor in a subsequent GLM (the PPI-GLM) together with a regressor for the experimental condition of interest (non-SV vs. SV) and a regressor modelling the interaction between the VOI and the experimental condition. The interaction between the VOI activity and the experimental condition was then used as an explanatory variable predicting neural activation in other brain regions. The regions experiencing a significant effect of the interaction regressor on their BOLD activity in the one-sample t-test group-level analysis were the ones affected by the region of interest in the context of non-SV vs SV in the whole sample.

3. 4. 3. Multivariate pattern analysis

In addition to the classical mass-univariate analysis, a multivariate pattern analysis (MVPA) method known as multivariate decoding was implemented to confirm that the different sacred values were perceived as semantically different categories by participants and to corroborate the univariate results. Unlike univariate methods, which predict one single dependent variable (i.e. the BOLD activation signal) for every voxel in the brain, MVPA methods simultaneously predict multiple dependent variables (i.e. BOLD activity in multiple voxels) as a function of one or more independent variables (Hebart, Görgen &

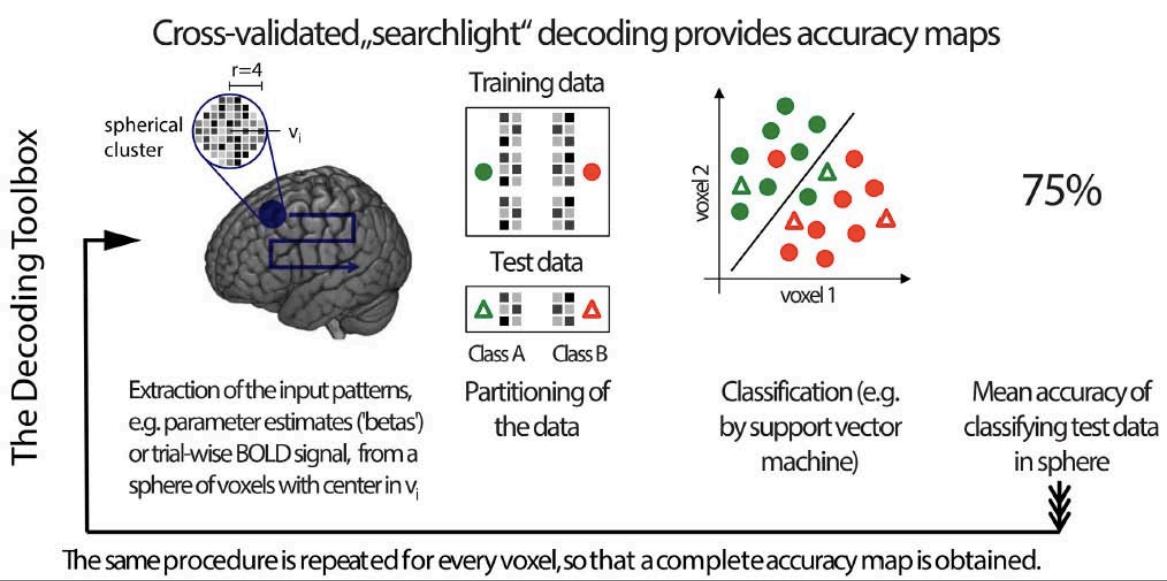
Haynes, 2015). The fact that activity in multiple voxels is analyzed conjointly allows accounting for their covariance, thus increasing the odds of finding statistical dependency (higher sensitivity). In addition, MVPA methods have a greater power to specifically detect information widely distributed throughout the brain in a fine-grained manner, e.g. differentiating individual faces (Kriegeskorte, Formisano, Sorger, & Goebel, 2007).

Box 4. Multivariate decoding analysis

The key concept in a decoding analysis is that information is represented in a hyper-dimensional vector space, where each dimension represents BOLD activity in a given voxel at a given time point. In a whole-brain decoding analysis, brain states or brain activity patterns are stored as vectors in this hyper-dimensional space, with each element of the vector representing the amount of BOLD activity found along a given dimension (i.e. voxel in time). Thus, whilst similar brain states (e.g. brain watching different human faces) are stored in nearby locations, more dissimilar brain states (e.g. brain watching a human face vs. watching an inanimate object) are located further apart. Making use of this information, a machine learning algorithm - in this case, a support vector machine (SVM) - is trained to classify distinct previously defined categories within a training set, i.e. a sample of brain states associated with categorical labels provided by the researcher. In terms of the hyper-dimensional space, the classifier learns to distinguish between categories by building a hyper-plane that works as a boundary between brain states from different categories. Subsequently, the trained classifier is tested to classify a different sample of unlabeled brain states. As a result, an above-chance classification accuracy rate ranging between 50% (classification at chance level) and 100% (all classifications correct) is extracted for the whole brain. Such accuracy rate is a measure of how distinguishable or classifiable two brain states are.

In the present study, a specific decoding modality known as searchlight analysis was performed. Instead of classifying whole brain states, this method employs a spherical cluster (in this case of radius 12mm) multiple times throughout the brain (see Figure 5). Thus, a smaller n-dimensional vector space, with n corresponding to the number of voxels in the cluster, is generated. Using this technique, the classifier learns to distinguish between different cluster states, or patterns of voxel activity within the cluster, resulting in

an above-chance classification accuracy rate for the sphere in a given brain location. This procedure is used throughout the whole brain in a serial manner, so that each time the sphere is centered on the next voxel in the brain. As a product, an above-chance accuracy map is obtained, including classification accuracies for each voxel in the brain, where the sphere was centered. Unlike the whole brain approach, the searchlight analysis allows to distinguish brain clusters that contain information for the distinction between categories above an established statistical threshold.



.

Figure 5. Multivariate decoding analysis by means of The Decodging Toolbox (TDT, Hebart et al., 2015). A spherical cluster centered in a given voxel in the brain is used to compare patterns of voxel activations between different classes of stimuli with a known label ("Training data set") and then tested in another data subset ("Testing data set"). Classification of stimuli is run by a support vector machine and results in an above-chance classification accuracy rate for the sphere in that given location. The sphere then shifts into the next voxel and the process is repeated, thus obtaining a mean classification accuracy rate for each voxel in the brain. The whole process can be systematically repeated using different Training and Testing subsets from the same data set in order to increase statistical power (cross-validation). Image adapted from Hebart et al. (2015) with permission of Frontiers in neuroinformatics.

Here, the Decoding Toolbox (TDT) (Herbart et al., 2015) was used in SPM12 to conduct a searchlight decoding analysis (see Box 4) to identify brain clusters containing relevant information to distinguish between i) three randomly chosen sacred values for each participant ii) sacred and non-sacred values, iii) peers-lower feedback and peers-agree

feedback for sacred values, iv) peers-lower feedback for sacred and non-sacred values. In addition, a regression analysis with self-reported moral outrage as predictor of informative neural regions in the classification between i) peers-lower feedback and peers-agree feedback for sacred values, and ii) peers-lower feedback for sacred and non-sacred values was conducted.

The present decoding analysis was conducted on non-smoothed brain images and a required a specific GLM (explained below). In order to increase statistical power, and given the structure of the present data (e.g. 2 conditions of interest with around 20 repetitions in each of the 2 runs), a leave-one-out cross-validation approach was implemented. Hence, data from each run was split in two chunks, resulting in 4 chunks that contained an approximately equal amount of trials from each condition of interest (e.g. SV vs non-SV). Decoding proceeded therefore in 4 steps. In each decoding step, the classifier used a given chunk as the testing set (unlabeled) and the rest of the chunks as the training set (labelled), taking each different chunk as the testing set exactly once. For this reason, the first-level GLM for the MVPA included 8 different regressors, one per condition (e.g. SV vs non-SV) and chunk. Due to the presence of non-responded trials, the number of trials per condition and chunk was not always exactly the same (unbalanced data). For a higher temporal resolution of the neural activity generated after each event, the BOLD response was convolved with a discrete finite impulse response function (fir) of length 16 and order 8. Thus, neural activity taking place within the span of 16 seconds after the event was decoded every 2 seconds, adding up to a total of eight "bins". As for the group-level analysis, a repeated measures ANOVA with bin number as within-subject factor (8 dependent levels) was conducted. The presented results include the effect of bin 3, which covers neural activity ranging from second 4 to 6 after the event's onset, coinciding with the hemodynamic response function (hrf) peak of activation used in the univariate analysis (5 seconds after the event).

3. 4. 4. Same analysis approach

The same analysis approach (Görgen et al., in preparation) is a systematic method that can be used before or after the MVPA analysis to detect possible confounds arising from the interaction between the experimental design and the use of MVPA. For instance, higher number of trials for condition A would involve more stable beta estimates for this

condition, thus biasing classification towards condition A. Similarly, trial order could be affecting decoding if the training set included “A followed by B” and “B followed by A” an unequal amount of times.

In order to determine if potential confounds that might be affecting MVPA classification, the SAA takes each of the measures under suspicion as input to classify between conditions (e.g. *SV vs. non-SV*) instead of the neural data associated with them. If the machine-learning algorithm gets to classify both conditions based on the proposed measures with a significant above-chance accuracy, these are most likely affecting MVPA results.

Here, three potential confounds were tested: reaction times (“RT”), willingness to fight and die ratings (“rating”) and trial order (“trial”). Additionally, in order to detect confounds due to the general data structure, a fourth variable containing random numbers (“rand”) was used as classification measure. Significant accuracy in the classification between sacred and non-sacred values based on any of these behavioral or random measures would be indicative of a confound.

4. Results

Experiment 1: Willingness to fight and die for sacred values

4. 1. Behavioral results

During the Rating1 paradigm, participants conveyed their willingness to fight and die significantly faster in the sacred value (SV) condition (mean = 4.93 seconds, sd = 1.23) compared to the non-sacred value (non-SV) condition (mean = 5.71 seconds, sd = 1.26) (Wilk's Lambda = .508, F(1, 29) = 28.07, p < .001) as illustrated in Figure 6. As expected, willingness to fight and die ratings were significantly higher for SVs (mean = 6.59 out of 7 points, sd = 0.49) than for non-SVs (mean = 4.76 out of 7 points, sd = 1.30) (Wilk's Lambda = .319, F(1, 29) = 62.00, p < .001).

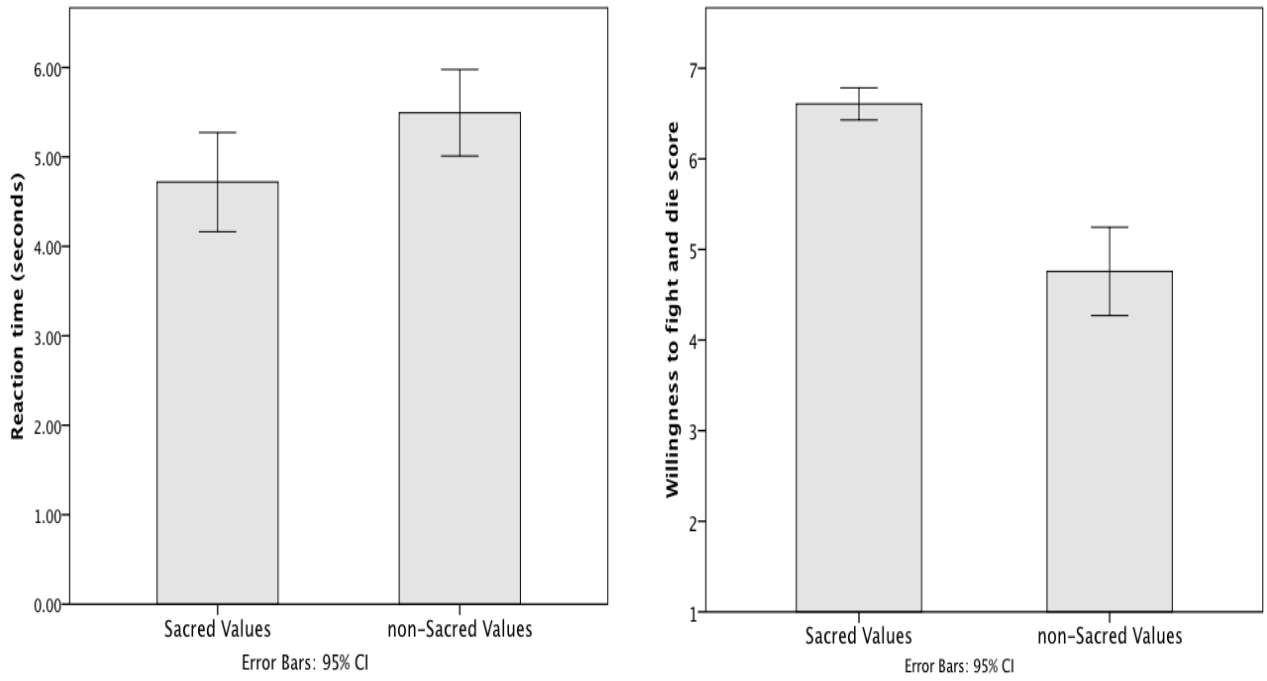


Figure 6. Behavioral results of the Rating 1 paradigm. On the left, reaction times were significantly shorter in the sacred value condition than in the non-sacred condition. On the right, willingness to fight and die scores (min = 1 and max = 7 points) were significantly higher and less variable for sacred values than for non-sacred values.

The item “Carry out militant actions to fight and die on my own, even if I had no group or support network” in the costly sacrifices questionnaire (see Annex, CS7) was highly correlated with the fifth (“Support in spirit or financially a non-state militant group”, r = .560, p = .001) and the sixth (“Join a non-state militant group”. r = .795 and p < .001) item of the same questionnaire. In addition, CS7 correlated positively with mosque attendance (r = .510, p = .004) and negatively with age (r = - 0.415, p = 0.023) and education level (r = -.383, p = .037). Indeed, a stepwise multiple linear regression analysis including these demographic variables as independent factors revealed a significant effect of the three factors on CS7 ($F(3, 26) = 8.17$, $p < .001$ with an $R^2 = .485$), with participants predicted CS7 = 6.27 + 1.46*Mosque attendance -.160*Age - .865*Education.

Moreover, CS7 positively correlated with preference for a more conservative European government (r = .456, p = .011), favorable opinion of LTJD (r = .382, p = .037) and the opinions “There is a war against Islam by Western nations” (r = .449, p = .013), “It really makes me angry when others criticize Muslims” (r = .536, p = .002), “I will never be satisfied until Muslims get the recognition they deserve” (r= .471, p = .009), “Muslims must unite in order to oppose Western forces by all means necessary” (r = .437, p = .016), “The strictest forms of sharia should apply to both Muslims and non-Muslims” (r = .544, p = .002), “Sharia should play a much larger role in Spanish law” (r = .483, p = .007). All items were measured by means of a Likert scale from 1 (“Strongly disagree”) to 5 (“Strongly agree”), means and standard deviation are presented in Table 1. A stepwise multiple linear regression including these opinions as independent variables revealed preference for a more conservative European government and “Muslims must unite in order to oppose Western forces by all means necessary” as significant predictors of CS7 ($F(2, 25)= 13.26$, $p < .001$ with an $R^2 = .515$), with CS7 = - 5.43 + 1.252*More conservative European Gov. + 1.397*“Muslims must unite in order to oppose Western forces by all means necessary”.

Table 1. Mean ratings and standard deviation of field survey items correlating with "willingness to carry out militant actions to fight and die on my own."

Field survey item (range: 1 to 5 points)	mean	sd
More conservative European government	2.10	1.03
Favorable opinion of LTJD	4.17	.379
"There is a war against Islam by Western nations"	4.53	.681
"It really makes me angry when others criticize Muslims"	4.13	.860
"I will never be satisfied until Muslims get the recognition they deserve"	4.47	.571
"Muslims must unite in order to oppose Western forces by all means necessary"	4.37	.490
"The strictest forms of sharia should apply to both Muslims and non-Muslims"	2.23	1.31
"Sharia should play a much larger role in Spanish law"	2.53	1.41

4. 2. Imaging results

4. 2. 1. Univariate analysis results

The main findings of the univariate analysis are shown in Table 2. The sacred value (SV) condition compared to baseline activity (fixation cross) was associated with increased activity in right temporoparietal junction (TPJ), right cuneus and a cluster including the right ventral anterior cingulate cortex and ventromedial prefrontal cortex (vmPFC/vACC) ($T = 4.97$, $p < .05$ FWE) (see Figure 7, in red-yellow). On the other side, the non-sacred value (non-SV) condition compared to baseline was also associated with right temporoparietal areas and a cluster comprising the right dorsomedial, inferior and orbitofrontal gyrus ($T = 5.00$, $p < .05$ FWE) (see Figure 7, in blue-green), but not vACC.

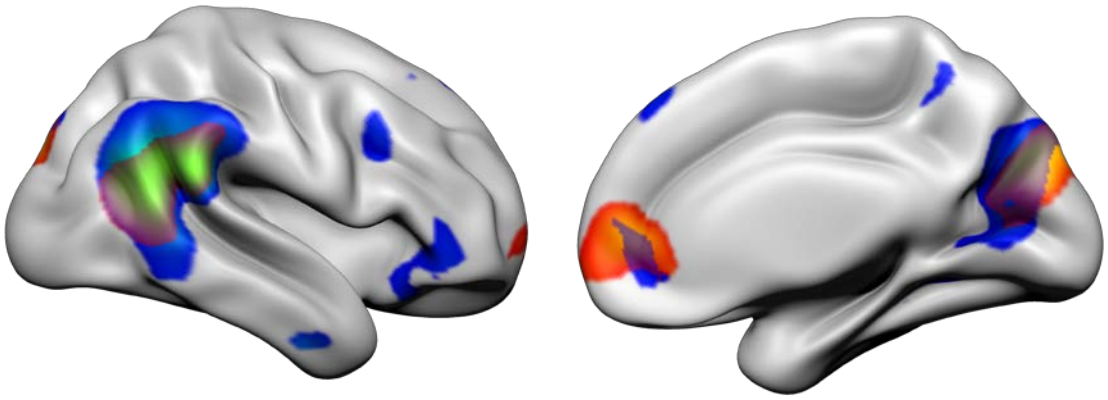


Figure 7. Results of the one-sample t-test analyses of Rating 1. Sacred values (SV, in red-yellow) and non-sacred values (non-SVs, in blue-green) showed similar patterns of neural activation when compared to baseline (fixation cross), including activity in the temporoparietal junction and the cuneus (T = 4.97, p<.05 FWE). However, prominent activity in the vACC/vmPFC was observed in SVs in contrast to non-SV against baseline (T = 4.97, p<.05 FWE).

The contrast between sacred and non-sacred values revealed decreased activity (see Figure 8 in blue-green) in the left dorsolateral prefrontal (dlPFC), left dorsomedial prefrontal (dmPFC), left superior parietal cortex and cerebellum (T = 5.06, p<0.05 FWE) and higher activity (see Figure 8 in red-yellow) in the bilateral cuneus (p<0.001 unc).

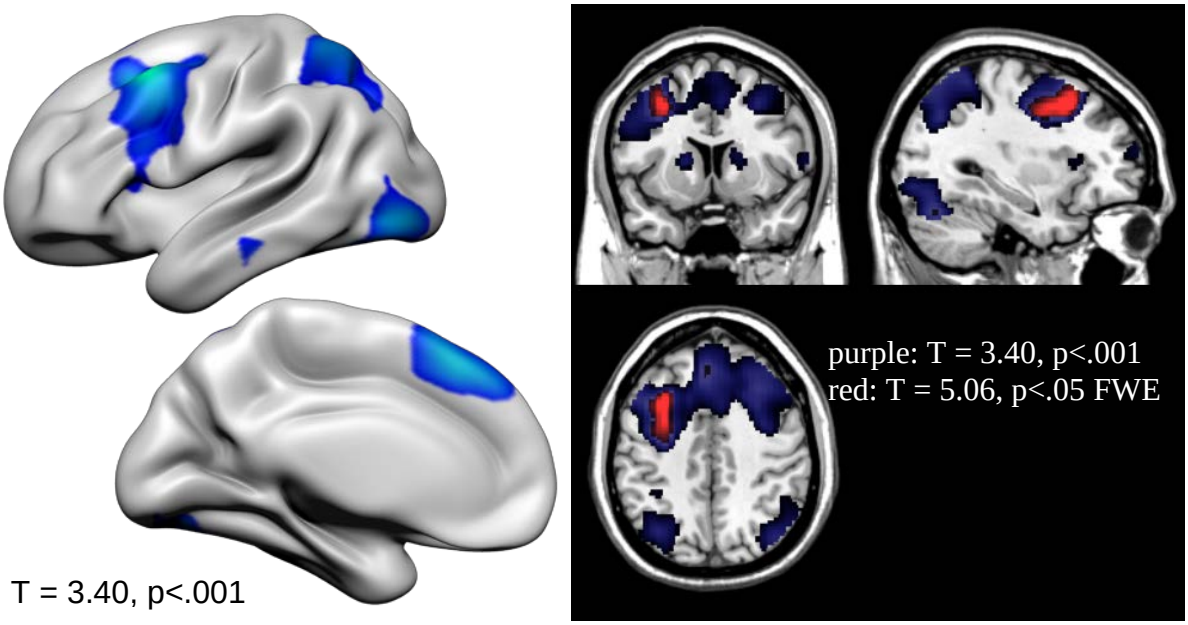


Figure 8. The contrast between SVs and non-SVs revealed decreased activity in the dorsolateral and dorsomedial prefrontal cortex, superior parietal cortex and cerebellum (blue-green) in the SV condition. On the left, surface projection of the resulting second-level T-map thresholded at p < .001 (T = 3.40). On the right, slice projection of the same T-map thresholded at p < .001 (T = 3.40) in purple and at p < .05 FWE in red.

The second-level regression analysis with reported spiritual formidability of the US (mean = 2.27 out of 5 points, sd = 1.11) as predictor revealed that participants who considered the US to be spiritually stronger exhibited higher activity in the left dorsolateral and superior medial prefrontal cortex while evaluating willingness to fight and die for both SVs and non-SVs against baseline (T = 5.09, p < .05 FWE) (see Figure 9).

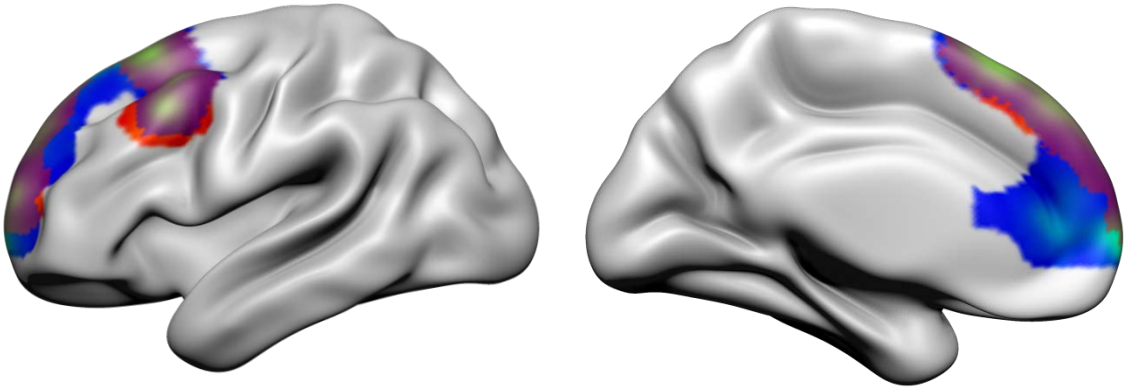


Figure 9. Results of the regression analysis including reported spiritual formidability of the US as predictor of neural activity during both SVs (red-yellow) and non-SVs (blue-green) compared to baseline. Strong overlapping can be observed for both conditions, with participants considering the US high in spiritual formidability showing greater activity in dlPFC, dmPFC, inferior parietal and cingulate regions during evaluation of SV and non-SV compared to baseline (T = 3.42, p < .05 FWEc, single voxel p < .001).

Table 2. Results of the Rating 1 univariate analysis including neural clusters surviving a p < .05 FWE correction across the whole-brain for the contrasts: sacred values > fixation cross, non-sacred values > fixation cross, non-sacred values > sacred values and US spiritual formidability predicting activity in non-sacred values > fixation cross and sacred values > fixation cross. The MNI coordinates of the local maxima peak, number of voxels, the highest Z-score and the cluster-level p-value after whole-brain FWE-correction are shown for each cluster. P-values marked with * were FWE-cluster corrected.

Univariate analysis: Main findings	Peak MNI coordinates			*N* of voxels	highest Z-score	FWE-corrected cluster-level p-value
	x	y	z			
Sacred values > fixation cross						
R middle temporal/supramarginal	58	-60	16	4433	6.30	<0.001
R anterior cingulate/ventromedial prefrontal	6	46	4	1376	5.51	<0.001
L lingual	-28	-52	-2	192	4.89	0.003
L middle temporal	-42	-70	18	143	4.76	0.005
Non-sacred values > fixation cross						
R angular/supramarginal	50	-56	26	9279	6.68	<0.001
L lingual	-28	-52	0	280	5.19	0.001
R inferior frontal opercular	40	16	32	299	4.98	0.001
R dorsomedial prefrontal	8	46	0	552	4.73	<0.001
R inferior frontal orbital/triangularis	32	32	-4	223	4.69	0.002
R dorsomedial/dorsolateral prefrontal	4	40	50	364	4.66	0.001
bilat precuneus	8	-52	58	244	4.50	0.002
Non-sacred values > Sacred values						
L dorsolateral prefrontal	-32	16	46	691	5.61	<0.001
R cerebellum	24	-78	-22	297	4.97	0.001
L dorsomedial prefrontal	0	26	52	278	4.74	0.001
L inferior frontal opercularis	-54	20	34	54	4.59	0.013
L cerebellum	-38	-72	-28	54	4.58	0.013
L superior parietal	-26	-68	62	2631	4.27	<0.001*
R superior parietal	44	-46	64	2565	4.16	<0.001*
US spiritual formidability predicting non-SV > fixation cross						
L dorsomedial prefrontal	-6	30	58	38	4.59	0.017
US spiritual formidability predicting SV > fixation cross						
L dorsolateral/dorsomedial prefrontal	-24	66	14	1517	4.21	0.007*

4. 2. 2. Effective connectivity analysis results

The effective connectivity analyses resulted in various observations (see Table 3). The vmPFC/vACC and TPJ, previously associated with sacred values compared to baseline (see univariate analysis results, section 4.2.1.), were used as seed regions in order to determine their respective contribution to further neural activity taking place in the same contrast (see Figure 10 left). As a result, in the context of sacred value assessment, the vmPFC/vACC exhibited a positive contribution to activity in a cluster comprising the left

posterior cingulate cortex and the precuneus up to the cuneus (T = 1.70, p < .05 unc) and a negative contribution to a cluster including the right dlPFC and the inferior frontal gyrus pars triangularis (T = 3.40, p < .05 FWEc, single voxel p < .001).

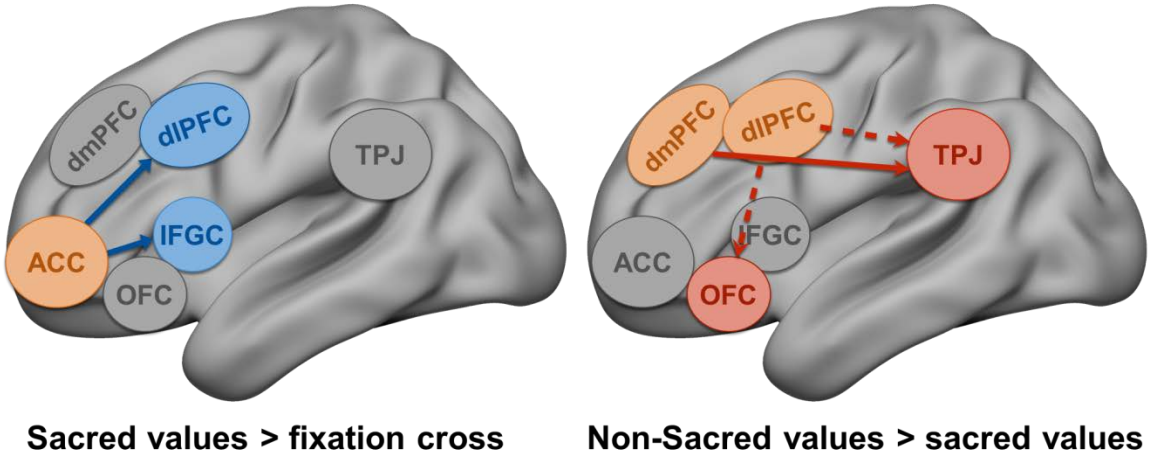


Figure 10. Effective connectivity model in the context of sacred values (left) and non-sacred values versus sacred values (right). In the SV condition, the vmPFC/vACC negatively affected activity in cognitive control structures such as the dlPFC and IFG. In the non-SV condition, the dmPFC and dlPFC positively contribute to TPJ activity. In turn, the dlPFC positively contributes to the OFC, an area associated with error prediction and evaluation. Seed regions are painted in orange, positive contribution arrows in red and negative contribution arrows in blue. Solid lines represent FWE-corrected contributions; dotted lines represent contributions thresholded at p < .001 uncorrected.

Additionally, the contribution of dlPFC and the dmPFC, which were significantly active during non-sacred value versus sacred value assessment (see univariate analysis results section), was assessed regarding further neural activity taking place in the same context (see Figure 10 right). Both structures revealed a similar pattern of contribution, positively influencing activity in the right TPJ (dmPFC: T = 4.99, p < .05 FWE; dlPFC: T = 3.40, p < .001 unc). Furthermore, the dlPFC also contributed to activity in the right orbitofrontal cortex (T = 3.40, p < .001 unc).

Table 3. Results of the Rating 1 connectivity analysis including neural clusters surviving a p < .05 FWE correction across the whole-brain for the positive interaction dmPFC x (non-sacred values > sacred values) and the negative interaction vmPFC/ACC x (sacred values > fixation-cross). The MNI coordinates of the local maxima peak, number of voxels, the highest Z-score and the cluster p-value after whole-brain FWE-correction are shown for each cluster. P-values marked with * were FWE-cluster corrected.

Connectivity analysis: Main findings	Peak MNI coordinates			*N* of voxels	highest Z-score	FWE-corrected cluster-level p-value
	x	y	z			
dmPFC (sphere centered at 0/26/52) postiviely contributing to contrast non-Sacred values > Sacred values						
R supramarginal	60	-40	32	143	4.56	0.005
vmPFC/ACC (sphere centered at 6/46/4) negatively contributing to contrast Sacred values > fixation cross						
R dorsolateral prefrontal/inferior frontal triangularis	50	24	40	744	4.09	0.040*

4. 2. 3. Multivariate pattern analysis results

The multivariate pattern analysis was congruent with the univariate analysis (see Table 4), confirming the role of the dmPFC, the dlPFC, the parietal cortex (left superior and inferior parietal lobe and right supramarginal gyrus) and the cerebellum as information-rich hubs for the distinction between sacred and non-sacred values (T = 5.2, p < .05 FWE). In addition, other relevant areas for an accurate classification between sacred and non-sacred values were detected, including the bilateral inferior frontal cortex, the right precuneus, the bilateral supplementary motor area and the middle cingulate cortex (T = 5.2, p < .05 FWE) (see Figure 11).

Table 4. Results of the Rating 1 multivariate analysis including clusters with differential activity surviving a p<.05 FWE correction across the whole-brain for the classification between sacred values vs non-sacred values. The MNI coordinates of the local maxima peak, number of voxels, the highest Z-score and the cluster p-value after whole-brain FWE-correction are shown for each cluster.

Multivariate analysis: Main findings	Peak MNI coordinates			*N* of voxels	highest Z-score	FWE-corrected cluster-level p-value
	x	y	z			
Peers-lower feedback (SVs) vs peers-agree feedback (SVs)						
L parietal operculum	-54	-16	22	101	4.16	0.002*
R cuneus	26	-60	22	106	4.13	0.002*
Moral outrage predicting Peers-lower feedback (SVs) vs peers-lower feedback (non-SV)						
L supramarginal/parietal operculum	-56	-22	18	59	4.35	0.028*

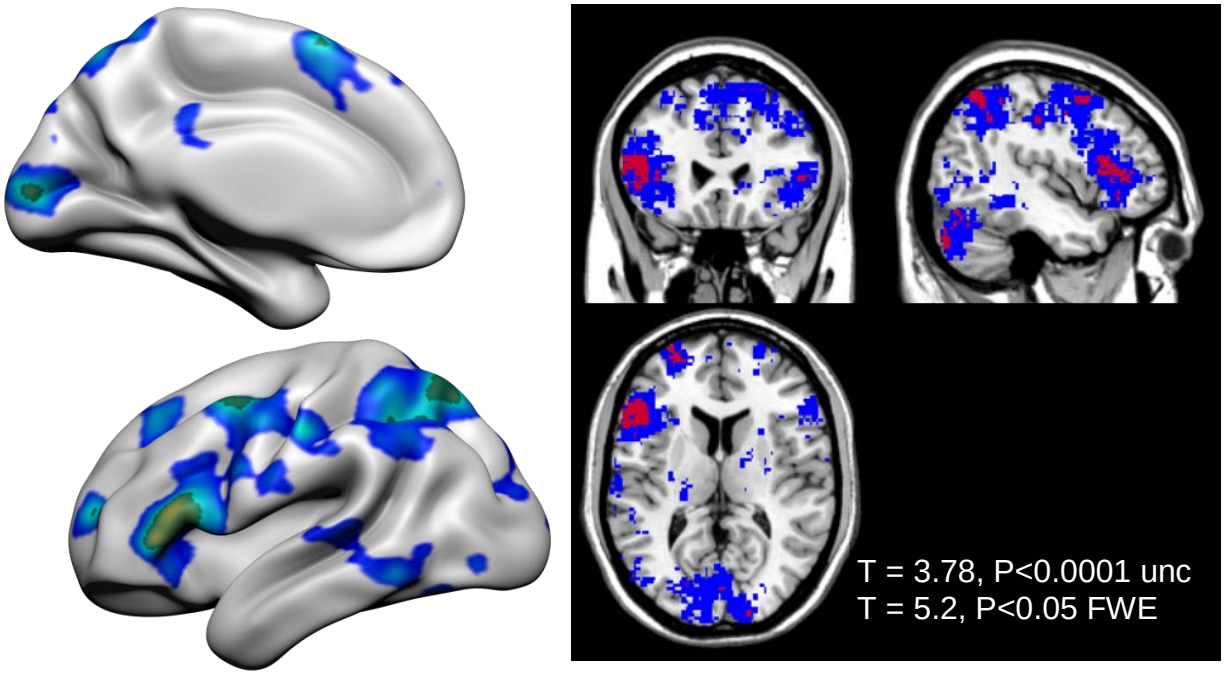


Figure 11. Results of a multivariate pattern analysis (MVPA) using support vector machine (SVM) classification between sacred and non-sacred values. On the left, surface projection of the resulting T-map thresholded at p < .0001 unc (T = 3.78) in blue and p < .05 FWE (T = 5.2) in red. On the right, transverse, sagittal and coronal slice projection of the same T-map thresholded at p < .0001 unc (T = 3.78) in blue and thresholded at p < .05 FWE (T= 5.2) in red. In line with the univariate analysis, the dmPFC, the dlPFC, the parietal cortex and the cerebellum revealed distinct patterns of activity between the sacred value and non-sacred value condition (T = 5.2, p < .05 FWE). Additionally, the increased sensitivity of the MVPA technique allowed to detect relevant differential activity in the bilateral inferior frontal cortex, the right precuneus, the bilateral supplementary motor area and the middle cingulate cortex between these two conditions (T = 5.2, p < .05 FWE).

Semantic differences between values

Multivoxel pattern analysis was also applied to conduct a classification between different sacred values, thus testing whether values could be distinguished as semantically distinct stimuli at a neural level. Three sacred values appearing at least 4 times during the paradigm in form of different grammatical rephrasings were randomly selected for each participant. Importantly, the set of three values was different for each participant. A linear support vector machine was then trained and tested to distinguish between these three semantic categories. The second-level analysis of the above-chance classification accuracy levels revealed a relevant role of the left middle temporal gyrus in making a semantic distinction between sacred values (T= 3.4, p<.001), regardless of which set of three specific sacred values was tested (see Figure 12).

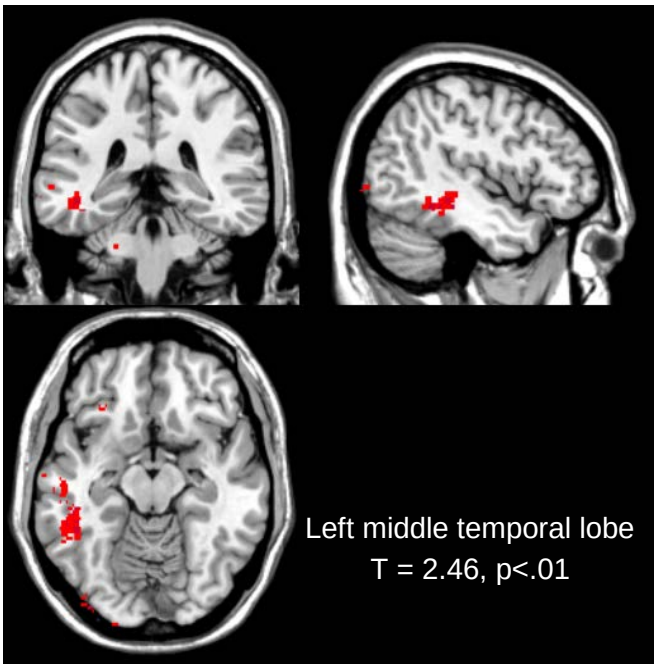


Figure 12. Slice projection of the resulting T-map (T = 2.46, p < .01) of an MVPA targeting neural activity patterns related to the semantic distinction between different sacred values. The left middle temporal gyrus consistently showed differential patterns of activity relevant to the classification between three randomly chosen sacred values (T = 3.40, p < .001).

4. 2. 4. Data quality check results

After checking the behavioral results, there were two very straightforward potential confounds to be assessed in the present data set. On one side, reaction times in the sacred value condition were significantly shorter than in the non-sacred value condition. Thus, a more stable beta estimate for non-sacred values could be biasing MVPA classification towards non-sacred values. On the other side, participants' willingness to fight and die ratings were significantly higher in the sacred value condition. Hence, the neural regions found to contain information relevant to the sacred vs non-sacred value distinction could also include information inherent to the behavioral responses, such as eye gaze towards the left (higher values) in the Likert scale for SVs, and the visual content related to it (higher numbers) (see Methods, section 4.4.4., for a more detailed explanation).

Due to such significant differences between SVs and non-SVs trials in reaction times and, especially, in willingness to fight and die ratings, these two behavioral measures were relatively useful for the SAA support vector machine (SVM) to classify between SVs and non-SVs (see Figure 13). Hence, these two measures most likely affected the MVPA

analysis based on neural data, and should be controlled for in future experiments. Other than that, and given that classification accuracy based on i) previous RT and rating and ii) both previous and current trial order and random numbers was either zero or relatively low, the general data structure did not look problematic.

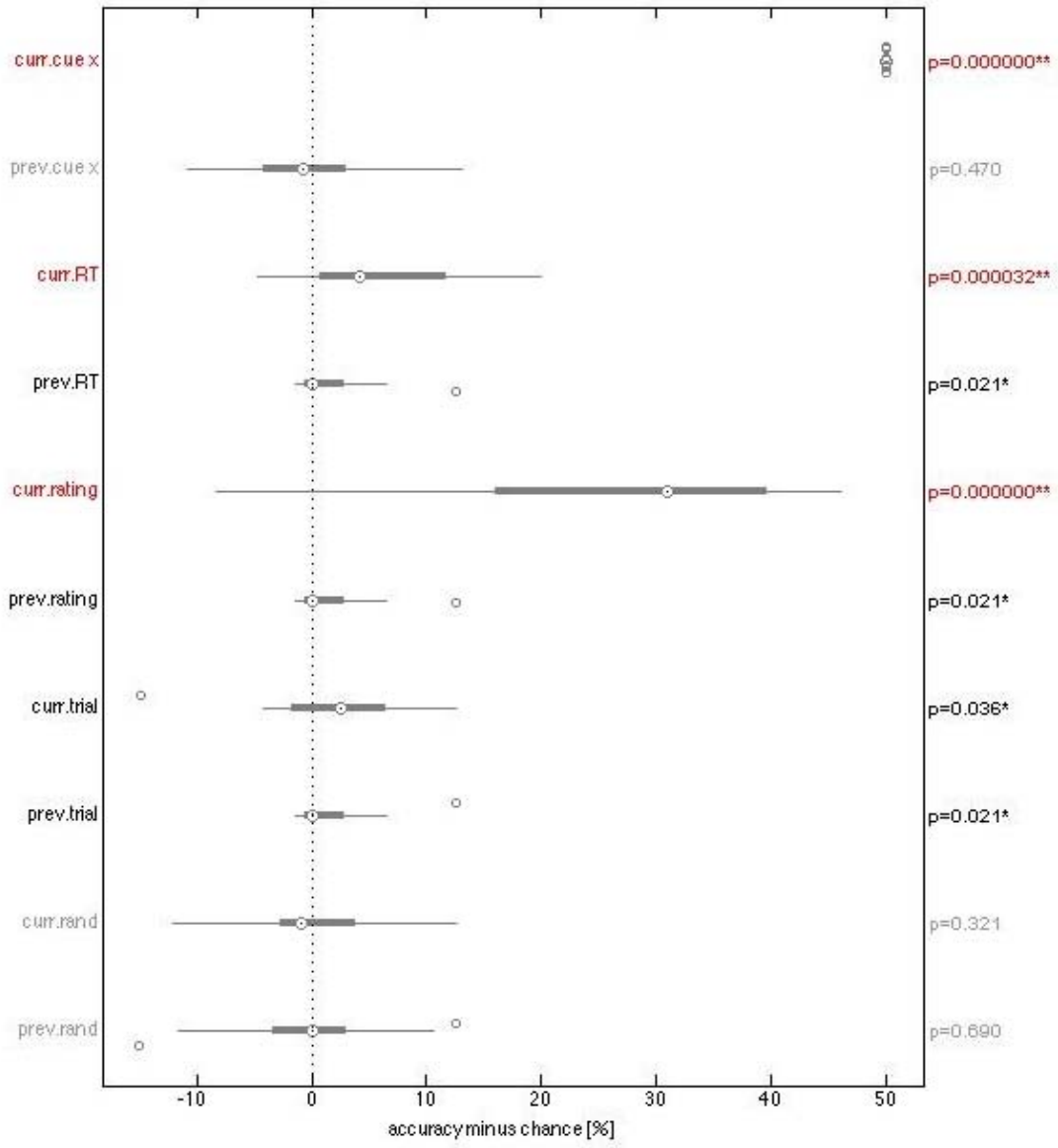


Figure 13. Outcome of the SAA analysis: Percentage of above-chance classification accuracy (x-axis) for each of the potential confounds (y-axis) including reaction time of the current ("curr.RT") and previous trial ("prev.RT"), willingness to fight and die ratings of the current ("curr.rating") and previous trial ("prev.rating"), trial number of the current ("curr.trial") and previous trial ("prev.trial"), and random number assigned to the current ("curr.rand") and previous trial ("prev.rand"). "Cue" is binomial and represents the condition of interest (either sacred or non-sacred value). Thus, classification accuracy above chance based on "Cue" values should be maximal (50%), as shown in the graph.

Experiment 2: Influence of social feedback on intentional action

4. 1. Behavioral results

In the feedback section, manipulated community ratings were calculated based on each participant's responses during Rating 1, and were programmed to create a *peers-higher* condition one third of the times, *peers-agree* another third, and *peers-lower* the rest of the times, as explained above. However, given that most of the time participants scored close to seven (mean = 6.59 out of 7 points, sd = .49) for SVs, the *peers-higher* condition could only rarely be implemented for sacred values. In fact, out of the thirty participants, only four received peers-higher feedback in at least one SV. For this reason, the *peers-higher* condition had to be dropped from the behavioral and the neurofunctional analysis.

No reaction time differences between SVs and non-SVs were found in the feedback paradigm. A repeated measures ANOVA analysis with sacredness (SV/non-SV), feedback type (peers-lower/peers-agree) and time (rating1/rating2) as within-subject factors and willingness to fight and die ratings as the dependent variable revealed a significant effect of the interaction between feedback type and time on willingness to fight and die (Wilk's Lambda = .783, $F(1, 29) = 8.03$, $p = .008$), proving that the experimental manipulation had a significant effect on participants' willingness to fight and die (see Figure 14). However, although time and value sacredness had a significant effect as independent factors (Wilk's Lambda = .574, $F(1, 29) = 21.55$, $p < .001$ and Wilk's Lambda = .362, $F(1, 29) = 51.04$, $p < .001$, respectively), the interaction between time, feedback type and value sacredness was non-significant.

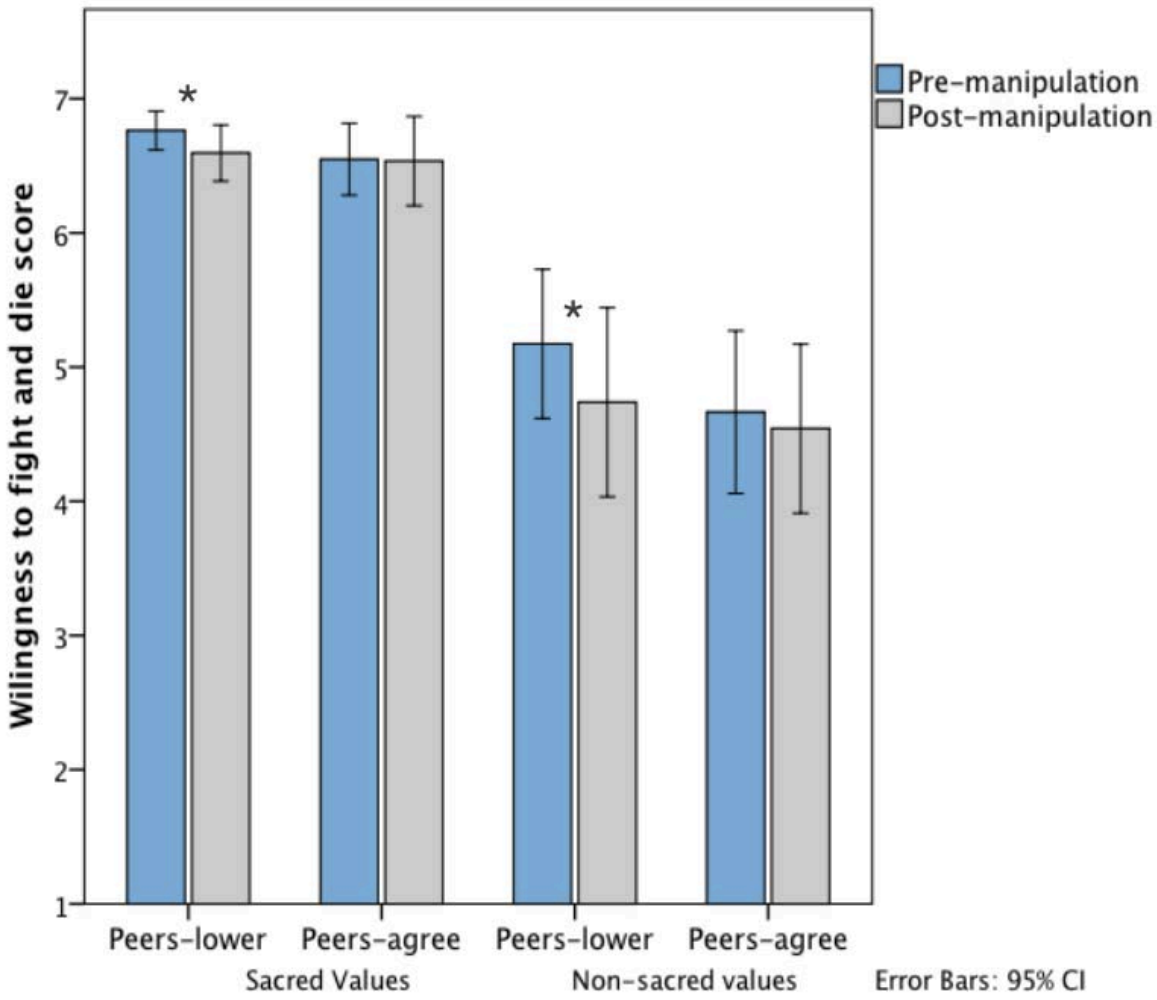


Figure 14. Willingness to fight and die ratings for sacred (left) and non-sacred values (right) before and after the community feedback manipulation represented as a function of feedback type (peers-lower/peers-agree). Post-manipulation ratings experience a drop in both sacredness conditions that is significantly steeper in the peers-lower feedback condition. The interaction between feedback type and time (pre/post-rating) had a significant effect on participants' ratings, with peers-lower feedback decreasing willingness to fight and die for both sacred and non-sacred values compared to *peers-agree* ratings. Although the drop between pre/post-ratings after peers-lower feedback was higher for non-Sacred Values than Sacred Values (.41 points vs. .25 points in a 7-point scale), the interaction between time, feedback and value sacredness was non-significant.

In addition, willingness to fight and die ratings for sacred and non-sacred values did not correlate with each other, higher sacred value ratings correlated with lower change in judgment (difference between *Rating1* and *Rating2*) for non-sacred values ($r = -.397$, $p = .03$), and resistance to change (inverse of change in judgment) for sacred values significantly correlated with higher change in judgment for non-sacred values ($r = .414$, $p < .023$).

The post-scanning emotions questionnaire revealed that participants experienced greater moral outrage in response to peers-lower feedback for sacred values (mean = 3.27 out of 7 points, sd = 1.93) compared to i) peers-agree feedback for sacred values (mean = 1.00, sd = .02, $t(29) = 6.44$, $p < .001$) and ii) peers-lower feedback for non-sacred values (mean = 1.84, sd = 1.13, $t(29) = 4.74$, $p < .001$). A repeated measures ANOVA with value sacredness (SV/non-SV) and feedback type (peers-lower/peers-agree) confirmed this effect, resulting in a significant interaction between value sacredness and feedback type on moral outrage ratings (Wilk's Lambda = .559, $F(1, 29) = 22.92$, $p < .001$) (see Figure 15 left). A similar result was observed for feelings of shame, with a significant interaction between value sacredness and feedback type (Wilk's Lambda = .567, $F(1, 29) = 22.18$, $p < .001$).

Joy ratings in response to peers-agree feedback for sacred values (mean = 5.60 out of 7 points, sd = 1.90) significantly higher than i) peers-lower feedback for sacred values (mean = 1.77, sd = 1.24, $t(29) = 10.02$, $p < .001$) and ii) peers-agree feedback for non-sacred values (mean = 4.24, sd = 1.86, $t(29) = 4.24$, $p < .001$). Again, a repeated measures ANOVA confirmed the interaction between value sacredness and feedback type (Wilk's Lambda = .721, $F(1, 29) = 11.19$, $p < .002$) (see Figure 15 right). Participants also reported higher feelings of compassion and pride in response to peers-agree feedback for sacred values compared to peers-lower feedback for sacred values and to peers-agree feedback for non-sacred values with a significant effect of the interaction between value sacredness and feedback type on compassion ratings (Wilk's Lambda = .707, $F(1, 29) = 11.99$, $p < .002$) and pride ratings (Wilk's Lambda = .616, $F(1, 29) = 18.09$, $p < .001$).

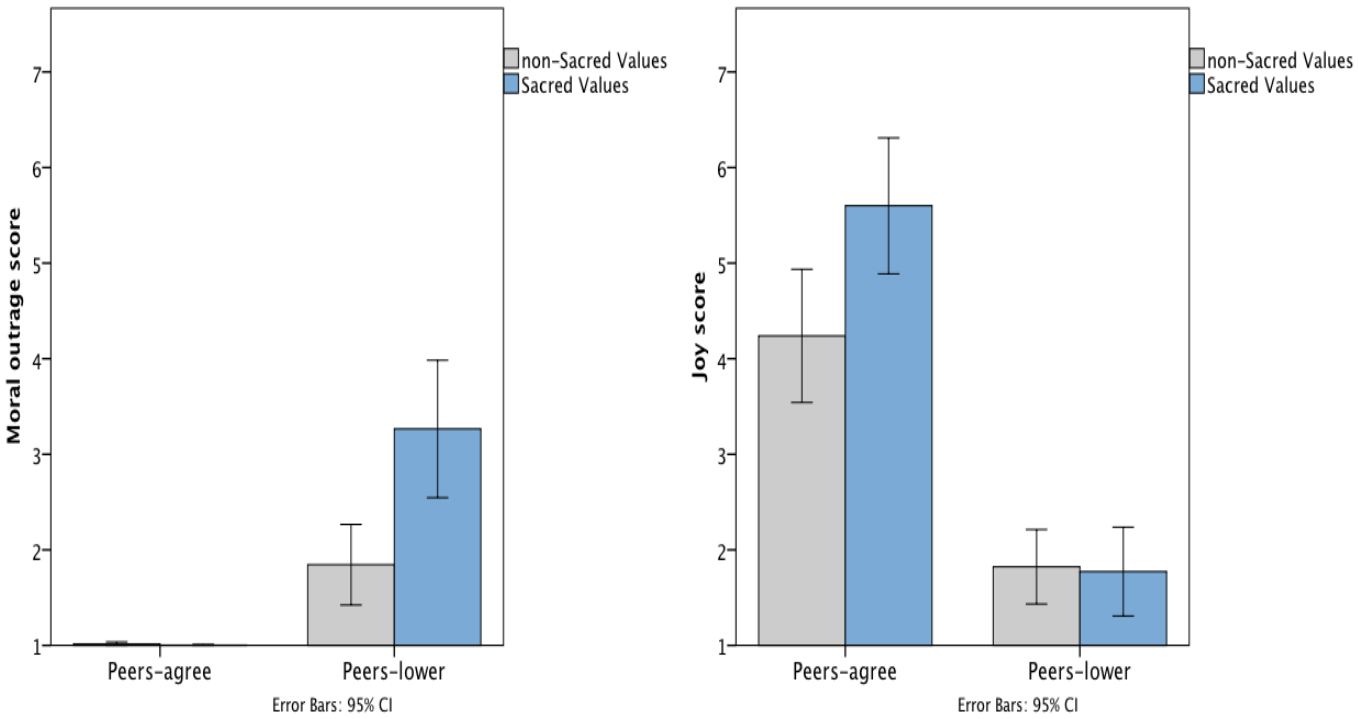


Figure 15. Significant effect of interaction between value sacredness and community feedback type on post-manipulation self-reported emotions towards the in-group. On the left, moral outrage as well as shame (not depicted) were significantly higher for sacred values in the peers-lower feedback condition than both for the peers-agree feedback condition and for non-sacred values. On the right, positive emotions such as compassion, pride and joy (depicted) were significantly higher for sacred values in the *peers-agree* condition compared to both the *peers-lower* condition and non-sacred values.

In fact, reported moral outrage in front of peers-lower feedback for SV positively correlated with reported shame for the same feedback condition ($r = 0.71$, $p < .001$) and with joy ($r = .55$, $p < .002$), compassion ($r = .63$, $p < .001$) and pride ($r = 0.57$, $p < .001$) in response to peers-agree feedback for SV. A more detailed description of the correlations between post-manipulation emotions is presented in Figure 16.

Moreover, the different post-manipulation emotions were evaluated with respect to sacred value ratings and change in willingness to fight after peers-lower feedback. Willingness to fight and die ratings for sacred values showed a tendency to correlate with outrage at peers-lower feedback to SV ($r = .349$, $p = .059$) and significantly correlated with shame at *peer-lower* feedback for non-SV ($r = .396$, $p = .037$). In turn, willingness to fight and die ratings for non-SV positively correlated with joy ($r = .369$, $p = .045$) and pride ($r = .361$, $p = .05$) at peers-agree feedback to non-SV. Resistance to change willingness to fight and

die for SV correlated with feelings of compassion (r = .387, p = .035) and pride (r = .378, p = .039) at peers-agree feedback to non-SV.

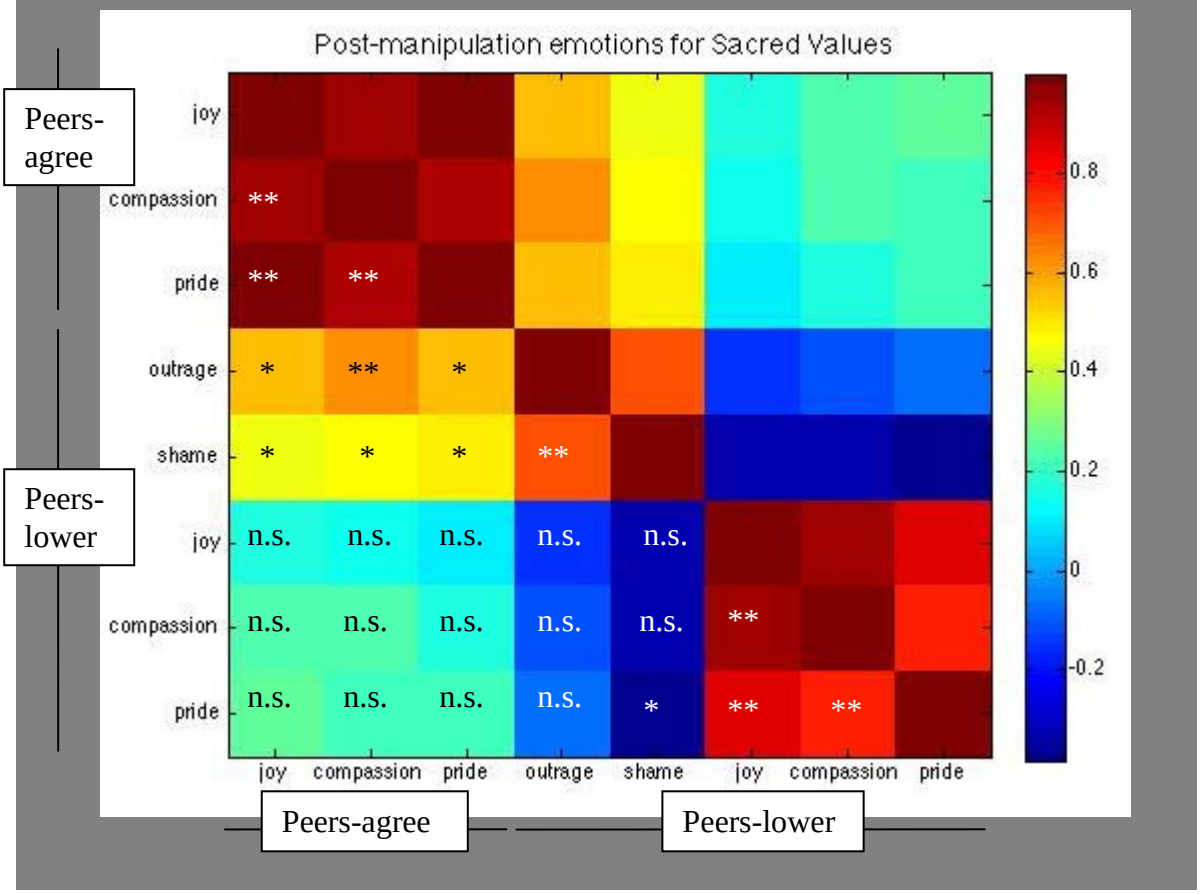


Figure 16. Correlation matrix of **post-manipulation emotions in the sacred value condition for *peers-agree* and *peers-lower* community feedback. Outrage and shame showed floor effects in the *peers-agree* condition and were thus not included in the correlation analysis. The color bar represents Pearson's coefficients r. Several observations could be made: a) High consistency among negative (outrage, shame) and among positive emotions (joy, compassion, pride) in all feedback conditions (red blocks). b) Outrage and shame towards the in-group after peers-lower feedback increased with positive emotions after peers-agree feedback conditions (orange and yellow blocks) c) Shame and pride after peers-lower feedback were negatively correlated (in blue). * p < 0.05, ** p < 0.001, n.s. = non-significant correlation.**

4. 2. Imaging results

4. 2. 1. Univariate analysis results

In the Feedback analysis, the peers-lower compared to the peers-agree condition for SVs (see Figure 17 left, in blue) was associated with activity in the right dlPFC, the left insula, and the left middle temporal region (T = 2.46, p < .01 unc). When compared to peers-lower feedback for non-SVs (see Figure 17 left, in red), the peer-lower condition for SVs exhibited higher activity in the bilateral insula, bilateral dlPFC and left superior parietal cortex (T = 2.46, p < .01 unc). However, these results did not survive multiple compassion correction. No significant activations were yielded by the interaction between feedback type and value sacredness.

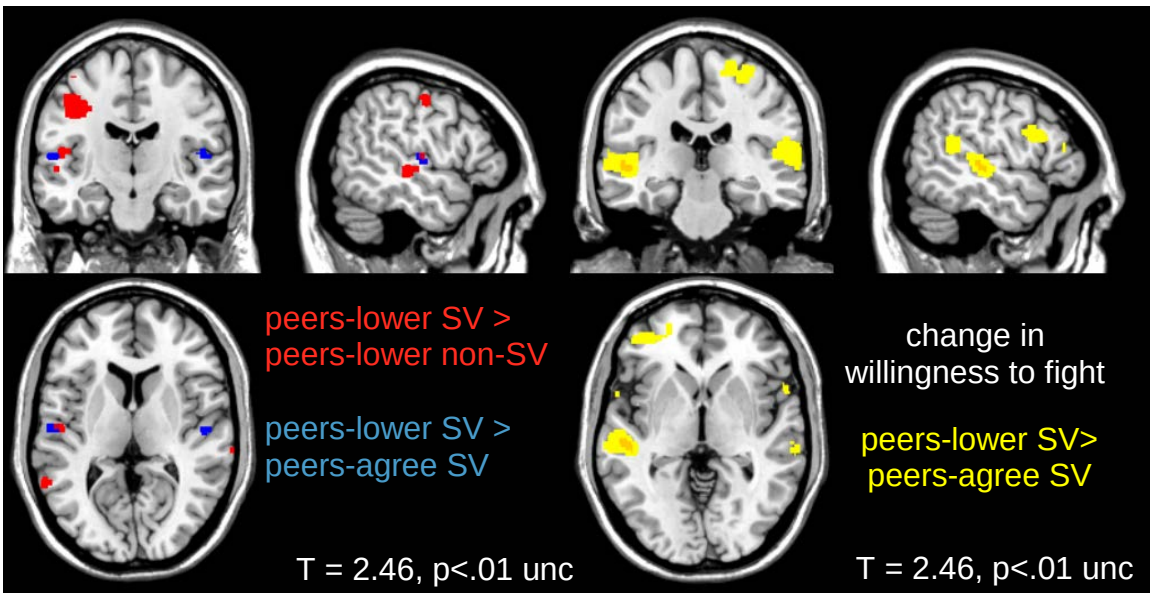


Figure 17. On the left, results of the one-sample t-test of the Feedback section. In blue, peers-lower feedback was associated with activity in the left insula and right dlPFC when compared to peers-agree condition for SV (T = 2.46, p<.01 unc). In red, peers-lower feedback for SV elicited higher activity in the bilateral insula, dlPFC and left superior parietal cortex in comparison to peers-lower feedback for non-SV (T = 2.46, p < .01 unc). On the right, activity in the left inferior frontal gyrus and insula while receiving *peers-lower* compared to peers-agree feedback was predicted by subsequent change in willingness to fight and die for sacred values (T = 3.42, p < .001, threshold used in slice projection: T = 2.47, p < .01).

Change in reported willingness to fight and die after receiving social feedback (difference between the intra-scanner and the post-scanning rating) was included as a regressor in a second-level analysis. Change in the SV condition successfully predicted activity in the

left inferior frontal opercular gyrus, left middle temporal and left insula while receiving *peers-lower* compared to peers-agree feedback for SV (T = 3.42, p <. 001 unc, see Figure 17 right).

4. 2. 1. Multivariate pattern analysis results

The MVPA analysis revealed differential activity in the left parietal operculum between the peers-lower and the peers-agree feedback condition for SVs (T = 3.12, p <. 05 FWEc, single voxel p < .001). Differential activity in the same area was successfully predicted by self-reported moral outrage at peers-lower feedback for SVs in the contrast between peers-lower feedback for SVs and non-SVs (T = 3.12, p < .05 FWEc, single voxel p <. 001). Results are summarized in Table 5 and illustrated in Figure 18.

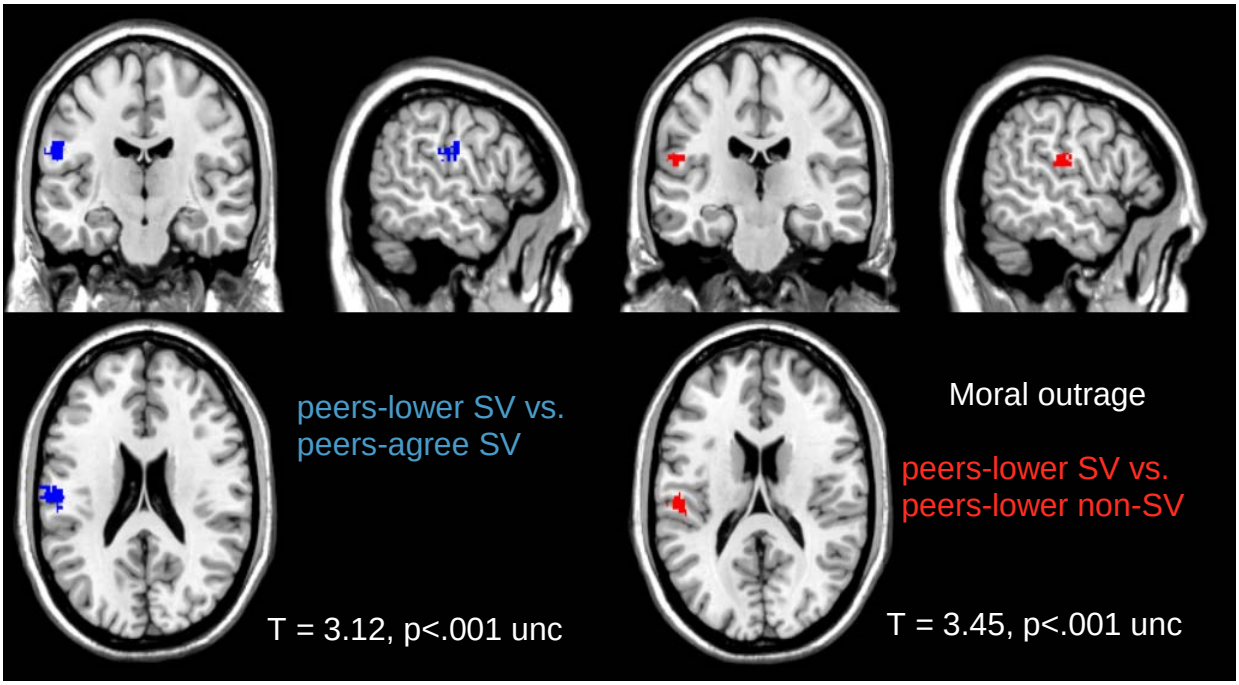


Figure 18. On the left, results of a one-sample T-test including above chance classification accuracy maps distinguishing between the peers-lower and peers-agree feedback conditions show differential activity of the left parietal operculum (projected values above T = 3.12, p < .001 unc). On the right, regression analysis results including self-reported moral outrage at peers-lower feedback for SVs as predictor of above chance classification accuracy between peers-lower feedback for SVs and non-SVs (projected values above T = 3.45, p < .001 unc).

Table 5. Results of the Feedback multivariate analysis including clusters with differential activity surviving a p < .05 FWE correction across the whole-brain for the classification between peers-lower feedback (SVs) vs peers-agree feedback (SVs), and activity differences in the classification peers-lower feedback (SVs) vs peers-lower feedback (non-SVs) predicted by moral outrage. The MNI coordinates of the local maxima peak, number of voxels, the highest Z-score and the cluster p-value after whole-brain FWE-correction are shown for each cluster. P-values marked with * were FWE-cluster corrected.

Multivariate analysis: Main findings	Peak MNI coordinates			*N* of voxels	highest Z-score	FWE-corrected cluster-level p-value
	x	y	z			
Peers-lower feedback (SVs) vs peers-agree feedback (SVs)						
L parietal operculum	-54	-16	22	101	4.16	0.002*
R cuneus	26	-60	22	106	4.13	0.002*
Moral outrage predicting Peers-lower feedback (SVs) vs peers-lower feedback (non-SV)						
L supramarginal/parietal operculum	-56	-22	18	59	4.35	0.028*

5. Discussion

Experiment 1: Willingness to fight and die in defense of sacred values

In the *Rating 1* paradigm the main findings include: i) higher and more consistent willingness to fight and die scores and lower reaction times for SVs than for non-SVs; ii) ventromedial prefrontal cortex/ventral anterior cingulate cortex (vmPFC/vACC) in SVs compared to baseline, and temporoparietal junction and cuneus involved in both SV and non-SV assessment,; iii) decreased dorsolateral (dlPFC) and dorsomedial (dmPFC) prefrontal, inferior frontal (IFG) and parietal activity during SV versus non-SV assessment; iv) individuals who reported higher spiritual formidability for the US exhibited higher activity in regions associated with non-SV processing (dlPFC, dmPFC and parietal) during both SV and non-SV assessment; v) in the SV context, the vmPFC/vACC contributed negatively to activity in cognitive control areas (dlPFC, IFG), which, in turn, positively influenced temporoparietal and OFC activity in the non-SV context.

5. 1. Behavioral observations: sacred values make one's own life look trivial

Behaviorally, the fact that willingness to fight and die (WFD) for sacred values (SV) was conveyed faster and was less variable between trials supports the notion proposed by Duc et al. (2014) of an adaptive role of SVs as an heuristic rule in decision-making. According to Duc et al.'s results (2014), trade-off situations involving sacred values versus mundane values (*taboo* trade-offs) were solved faster than trade-offs confronting two SVs (*tragic* trade-offs) in the group having more SVs. Under this framework, WFD for SVs can also be understood as a trade-off between a SV and one's own sacrifice. Remarkably, responses to this trade-off were conveyed faster than in the case of non-SVs, and overwhelmingly resolved in favor of SVs. This way, our participants prioritized the SVs in question over their biological self-interest. Conversely, trade-offs involving WFD in favor of non-SVs were more similar to Duc's *tragic* trade-offs, which required more time to respond, in a process that might have involved longer deliberation.

In addition, both Duc et al., (2014) *tragic* trade-offs and Li et al. (2011) *vital-vital* loss choices were associated with stress and negative affect, whereas *trivial-vital* choices correlated with positive affect in Li et al. (2011). Although we did not measure emotions during Rating 1, participants reported increased feelings of joy upon peers' agreement in WFD for SVs in comparison to non-SVs in the social influence experiment (discussed below). Thus, feelings of joy were not uniquely associated with peers' agreement in itself, but also with the idea of fighting and dying in defense of SVs in contrast to non-SVs. Such feelings of joy are in alignment with Li et al. reported positive affect in the *trivial-vital* choice. In other words, there must be a joy in renouncing to the mundane in support of vital or sacred causes. Note, however, that the notion of sacred choices in the present experiment surpasses that of Li's *vital* choices, since the trade-off proposed here involves giving up one's own life, turning what is vital into something comparably trivial. Thus, a hierarchical relationship encompassing the mundane, the vital and the sacred in ascending order could be established, being choices between unequal classes easier and faster to solve (a process possibly accompanied with feelings of joy), and choices between items of the same class stressful and difficult to ponder. Assuming that the sacred is in the utmost position, transcending the vital, it is not surprising to observe such an existential commitment to protect it. But how does commitment to the sacred look in the brain?

5. 2. Neurofunctional correlates of sacred and non-sacred values

Overlapping activity of the TPJ in WFD for both sacred and non-sacred values

The neurofunctional analysis revealed widely overlapping regions of activity between sacred and non-sacred values, including the right temporoparietal junction (TPJ). This is not surprising given that the presented items for both conditions were all Muslim cultural values, only differentiated by the aspect of sacredness to the particular participant (rejection of both financial concessions and democratic consensus). The TPJ has been typically involved in social cognition and moral decision making: whereas Berns et al. (2012) report increased TPJ function in decisions involving SVs, Greene et al. (2001) found TPJ was associated with personal compared to impersonal moral dilemmas, and Kahane (2012) with deontic compared to utilitarian reasoning. However, we found

increased TPJ activity both in the sacred and the non-sacred condition, which could indicate that both sacredness conditions involved a certain degree of morality, judging what would be "right or wrong", following Kahane's interpretation.

According to Greene et al. (2002), inducing personal harm could be conceptually atomized into "me hurt you", with the TPJ being responsible for computing the "you", i.e. there is an individual "different than me" receiving that harm. This idea is supported by other studies pointing at the role of TPJ in differentiating the self from the other (Decety & Lamm, 2007; Hogeveen et al., 2015), detecting movement and intentionality of other social agents (Borg, Hynes, Horn, & Grafton, 2006). Here, we asked for willingness to fight and die to defend either sacred or non-sacred values. Thus, it was to expect that the "you" aspect of inducing personal harm would be present in both conditions. In turn, the fact that Berns et al. (2012) did only find TPJ in the SV condition could be explained by the items that were used in the non-SV condition. Whereas we included cultural values that still contained a certain moral burden, without reaching the "sacredness status"(e.g. "women should wear hijab in public" was non-sacred for most of the sample), Berns et al. used quite trivial non-sacred values (e.g. "you are a dog person", "you are a Pepsi drinker"). Therefore, it is plausible that the non-SV condition in Berns et al. (2012) did not show specific TPJ activity. Here, we give proof that TPJ activity is not specific to SVs (as could be interpreted from Bern's findings), but rather, it is involved in moral judgment and perhaps, as Greene proposes, in encoding the concept of "other" as receptor of harm.

Prominent role of the vmPFC/vACC during sacred value assessment

Willingness to fight and die (WFD) for SVs but not non-SVs compared to baseline activity was very robustly associated with activity in the vmPFC/vACC. The vmPFC has been typically related to the affective aspects of decision-making in favor of an adaptive behavior (Koenigs et al., 2008), associated with empathetic responses (Farrow et al., 2001) and deontic social moral judgment (Young et al., 2007). It is worth noticing that the ACC is subdivided in two functionally distinct main nodes. On one side, the cognitive division (dACC), encompassing dorsal aspects of the ACC, has been associated with attentional modulation, action selection and conflict monitoring among other functions (see Bush, Luu, & Posner, 2000 for a review). On the other side, the affective division (vACC),

including the subgenual and pregenual portions (ventral aspects), is densely connected with the vmPFC and the amygdala and has been associated with affective appraisal (Etkin et al., 2011). In the present study, WFD for SVs was associated with a cluster of activity including both the vmPFC and the affective division of the ACC, which will be referred as vmPFC/vACC for simplicity (see Figure 19).

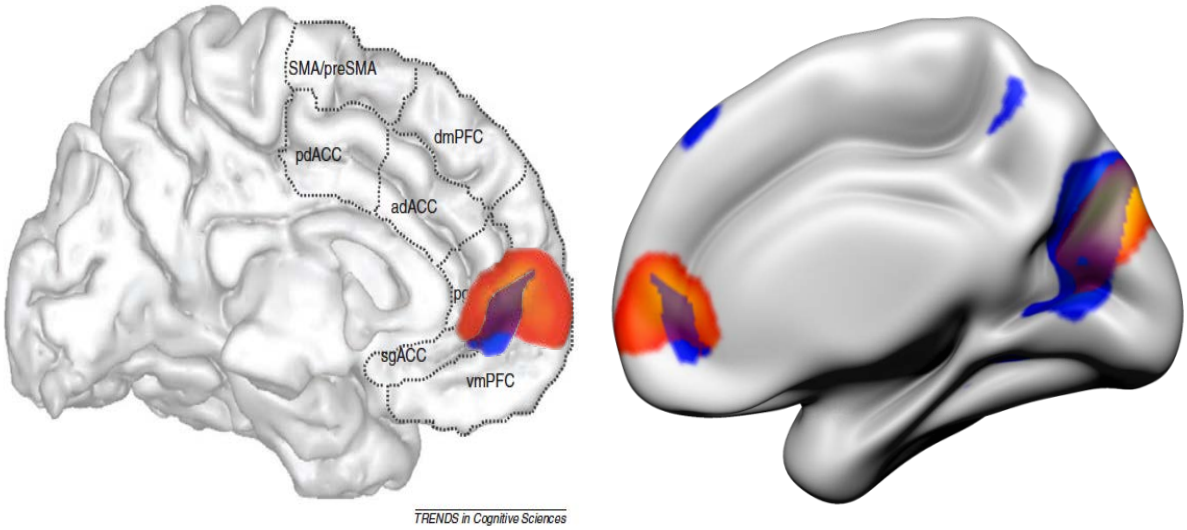


Figure 19. Subdivisions of the anterior cingulate cortex (ACC) and the medial prefrontal cortex (mPFC) with the projected results of the contrast SV and non-SV against baseline in red and blue, respectively. The dorsal aspects of the ACC encompass the posterior and anterior parts (pdACC and adACC), and the ventral ACC includes pregenual and subgenual portions (pgACC and sgACC). The medial prefrontal cortex is divided in more dorsal aspects (supplementary motor area/SMA and dorsomedial PFC/dmPFC) and ventral parts (ventromedial PFC/vmPFC). Illustration adapted from Etkin et al. (2010), reproduced with permission of Trends in Cognitive Science.

Many different studies point at the vmPFC/vACC as a key node for self-referential processing, i.e. processing of stimuli both in the sensorimotor and emotional domain that are experienced in close relation to one's own person, e.g. watching images of the self, close friends and family, or personal objects. For instance, Northoff et al. (2006) reviews extensive evidence supporting that anterior aspects of the cortical midline structures (CMS), including both the vmPFC and pregenual ACC, are polymodal convergence zones (Rolls, 2000) responsive to high degrees of self-referentiality (see especially D'Argembeau, 2013). Some of the reviewed studies refer specifically to the role of the vmPFC in first-person perspective processing (Ruby, Decety, Inserm, Thomas, & Cedex, 2003; Vogeley et al., 2004). In turn, Schmitz & Johnson (2007) propose the vmPFC-ACC-subcortical network as top-down regulatory system both selecting non-salient self-relevant

information and filtering out non-self-relevant sensory information. Importantly, the ventral CMS might not only encode self-referential processes referred to the self but also to those similar to the self (Mitchell et al., 2006). In accordance to this interpretation, higher vmPFC/vACC activity during SV assessment could be due to increased self-relevance of the values presented in the SV condition, possibly surpassing that of one's own life (see discussion on behavioral results, section 5.1.). Although the contrast between SVs and non-SVs does not show a significant increase in vmPFC/vACC for SVs, the contrast with baseline shows a strong activation in the SV condition that survives even at $P < .001$ FWE ($T = 6.75$), which is, however, not detected in the non-SV condition versus baseline, even at lower thresholds. In addition, the connectivity analysis highlights the importance of the vmPFC/vACC for SV processing, as it inhibits the dlPFC in the SV context, in line with the decreased dlPFC activity observed in the SV versus non-SV contrast.

Several other studies highlight the role of the vmPFC/vACC in volition and self-initiated action. According to Brass & Haggard (2008), the vmPFC might process stimulus-independent information from interoceptive centers (Buckner, Andrews-Hannah & Schacter, 2008) producing internally guided behavior, i.e. "voluntary" action (Passingham, Bengtsson, & Lau, 2009). More recently, Haynes, Wisniewski, Görgen & Momennejad (2015) detected mPFC activity in association with self-initiated or freely chosen intentions compared to cued intentions. Along these lines, Tanaka, Balleine, & Doherty (2008) found that the vmPFC responded to contingency between participants' actions and reward presentation, creating the subjective experience of causality or agency. Hence, vmPFC activity might underlie feelings of agency and perceived behavioral control. Interestingly, high levels of perceived control over one's behavior (measured as internal "locus of control", see Box 5) correlated with increased grey matter volumes in the ACC extending into the vmPFC in a voxel-based morphometry study including 777 young adults (Hashimoto et al., 2015). Moreover, Harnett et al., (2015) observed that individuals with poor perception of control exhibited decreased vmPFC activity in response to predictable threat.

In the light of these studies, the increased vmPFC/vACC activity observed in association with willingness to fight and die for SV could be related to a stronger perception of agency and control. Indeed, an increased personal control or sense of agency (SoA) through the

expression of WFD for SVs would comply with the predictions derived from the devoted actor model. In the first place, devoted actors are often part of groups such as liberation armies that arise in defense of those who they feel have been humiliated by a materially stronger foe (Atran, 2003), e.g. Western nations against the Muslim world, India against the Kashmiris or Israel against Palestinians. For these militants, the unequal distribution of power legitimizes the use of violence as the only way out against what they see as extreme acts of injustice. Hence, for devoted actors, the idea of fighting and dying for SVs might increase feelings of control and self-agency in a situation otherwise dominated by a materially stronger rival. Crucially, feelings of personal control or sense of agency (SoA) are intimately related to power (Fast, Gruenfeld, Sivanathan, & Galinsky, 2009; Inesi, Botti, Dubois, Rucker, & Galinsky, 2011; Obhi, Swiderski, & Brubacher, 2012). Thus, it seems likely that existential commitment to group-related SVs might, at least in part, compensate the loss of control and powerlessness induced by conditions of war or perceived abuse by a higher power.

In the second place, devoted actors are characterized by making decisions independently of prospects of success (Atran et al., 2007). For instance, despite of having an army several magnitude orders smaller than his opponent, Leonidas chose to fight the Thermopylae battle. In line with the devoted actor framework, his decision was probably not responding to the rational odds of winning, but rather to reasons of honor or to avoid humiliation. Indeed, a recent study gives evidence that perceived sense of control has a rewarding effect independent of outcomes that biases action selection in a way similar to tangible rewards (Karsh & Eitam, 2015). In it, participants preferably chose actions conferring the highest levels of agency regardless of the outcomes, highlighting the motivational properties of being in control. Sense of agency has been pointed out as an adaptive biological need (Leotti & Ochsner, 2011). Hence, personal search for control and sense of agency might be expressed through willingness to fight and die for SVs in conditions of loss of personal control or powerlessness that can otherwise lead to depression and anxiety (Calhoun, Cheney, & Dawes, 1974; Weems, Silverman, Rapee, & Pina, 2003). In fact, the ACC (activated at the thought of fighting and dying for SVs) has been found to be related with a positive self-bias (*superiority illusion*, Yamada et al., 2013), whereas reduced ACC volume has been associated with major depression and posttraumatic stress disorder (Herringa et al., 2012). In line with these observations, future experiments should include a measure of perceived sense of control and positive self-bias

that could potentially be useful to predict vmPFC/vACC activity while reporting WFD for SVs.

Box 5. Locus of control

Locus of control is a psychological construct pinpointed by Rotter et al. (1966) that represents the degree to which an individual attributes the causality of events to himself (internal locus of control) or to the environment (external locus of control). Whereas internal control has been associated with emotional stability, physical and psychological well-being and stress resilience (Bollini, Walker, Hamann, & Kestler, 2004; Judge & Bono, 2001), external control is related to lower tolerance to aversive events, neuroticism and even anxiety and depressive symptoms (Barlow, Ellard, Sauer-zavala, Bullis, & Carl, 2014; Koolhaas et al., 2011). However, a moderating effect of culture type on the correlation between external LOC and anxiety has been described (Cheng, Chio, & Chan, 2012), which is increased in individualistic societies that prioritize individual over social goals, and decreased in collectivistic ones, where social harmony is emphasized. Interestingly, the internal LOC functions as a mediator in the relationship between religious beliefs and psychological well-being (Ryan & Francis, 2012).

Finally, the increased vmPFC/vACC activity observed in the SV condition could also be explained by higher perceived self-trascendence when reporting WFD for SV. Hakamata, Iwase, Kato, Senda, & Inada (2013) found that self-trascendence measured with the Cloninger Personlity Test (Cloninger et al., 1993) successfully predicted resting brain glucose metabolism in the sgACC extending into the vmPFC. Self-transcendence is a poorly understood psychological construct that includes beholding spiritual/religious/supernatural beliefs, dissolution of the self and connection with other beings (MacDonald et al., 2002). In front of these results, it would be interesting to test whether, in conditions of humiliation and powerlessness, existential and "selfless" commitment to cultural or religious values might serve as a means to retrieve the lost sense of personal agency and self-transcendence. For now, following Greene et al. (2002) proposed "me hurt you" model of morality and given the vmPFC/vACC implications in self-agency, we are inclined to think that the increased vmPFC/vACC activity found in

association with WFD for SV is related to a heightened perception of the subject "me", thus boosting the sense of agency.

Decreased dorsal prefrontal activity during sacred value assessment

The contrast between WFD for SVs and non-SVs revealed a strong reduction in dorsolateral (dlPFC) and dorsomedial (dmPFC) prefrontal activity in the SV condition. These results support Greene et al. (2004) findings that associate deontic responses with decreased dlPFC activity and shorter reaction times in comparison to utilitarian decisions. In this sense, the question on WFD for sacred values, similarly to Greene's personal moral dilemmas, is resolved faster and with lower involvement of the dlPFC than WFD for non-sacred values. As presented in the introduction, the dlPFC has been involved in cognitive control, abstract reasoning (Koechlin, 2003; Ramnani & Owen, 2004) and reappraisal of emotional stimuli (Ochsner, Bunge, Gross, & Gabrieli, 2002; Ochsner & Gross, 2005). Hence, deciding upon non-SV elicited a higher recruitment of structures engaged in problem-solving and deliberation than when SVs were invoked.

In turn, the decreased dmPFC activity in the SV condition could be due to a limited ponderation between different behavioral outcomes. The dmPFC has been involved in reappraisal and explicit reasoning of self-relevant stimuli as part of the dorsal cortical medial structures (Northoff et al., 2006) and given its dense connections with the dlPFC (Öngür & Price, 2000). Moreover, Alexander & Brown (2011) argue that the dorsal aspects of the mPFC might encode different possible predicted results and respond to unexpected outcomes, allowing to learn new contingencies. According to these authors, the dmPFC might compute the value of anticipated events (Rudebeck et al., 2008) as a result of cost-benefit calculations (Croxson, Walton, Reilly, Behrens, & Rushworth, 2009). In turn, dmPFC has also been involved in action selection and "intentional stopping", i.e. deciding whether or not performing an action (Brass & Haggard, 2007). Thus, the reduced engagement of the dmPFC in SV assessment suggests a more modest elicitation of cost-benefit analysis and value computation of anticipated action. This interpretation is supported by the shorter reaction times found in the SV condition, that point to more automated responses (Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2009), likely bypassing a more scrupulous cost-benefit ponderation.

Intriguingly, Berns et al. (2012) did find an association between dmPFC activity and SV assessment that the authors attribute to a higher recruitment of theory of mind (ToM) circuitry in deontic reasoning. However, this result opposes ours, which is better sustained by previous moral judgment literature pointing at the dmPFC as an action selection node. In this case, Berns' result could be due to the way SVs were presented, in form of a second person statement ("You believe in God"). This sentence structure could indeed call upon mentalization processes, given the presence of a subject emitting a judgment in second person that participants understand as the self. If this is true, mentalization would be highly dependent on the sentence structure rather than to the sacredness of the presented value.

Other differences between SV and non-SV assessment

The univariate analysis results also include a decreased activation of the superior parietal cortex in the SV compared to the non-SV condition. This result cannot be directly assimilated to the involvement of the inferior parietal cortex in utilitarian decisions found by Greene et al. (2004) and in non-SV assessment observed by Berns (2012). On one side, the inferior parietal cortex has been involved in task-reward associations flexibly adapting to trial-by-trial changes in task-reward contingencies (Wisniewski, Reverberi, Momennejad, Kahnt, & Haynes, 2015). On the other side, Kahnt et al., (2014) analyzed the independent effects of salience and value computation in task-reward associations, and found that the inferior parietal cortex specifically responded to salience of predicted outcomes, whereas the superior parietal cortex encoded the predicted value of tasks. Thus, it is plausible that whereas Greene (2004) and Berns (2012) detected inferior parietal activity associated with outcome salience in the utilitarian/non-SV condition, the present data possibly shows superior parietal responsiveness to value attribution in front of task-reward contingencies invoking non-SVs, i.e. fighting and dying for a non-SV. Therefore, the increased superior parietal activity in non-SV trials associated with value computation within a framework of task-reward evaluation suggests that, similarly to Greene (2004) and Berns (2012), non-SV decisions are sensibly more reliant on utilitarian reasoning than SV decisions.

The multivariate analysis (MVPA) results confirmed the differential activity of the dlPFC, dmPFC and the parietal cortex in the SV and non-SV conditions. Interestingly, the MVPA detected differences in both the superior and inferior aspects of the parietal cortex. Both

superior and inferior parietal activations have been related to task-reward associations, although no inferior parietal activity differences were detected in the univariate analysis. In the MVPA analysis, inferior parietal activity could be more readily linked to predicted outcome salience (Kahnt et al, 2014), as explained above.

Additionally, the MVPA analysis was sensitive to differential activity in the inferior frontal gyrus (IFG) and the right precuneus, bilateral supplementary motor area, and middle cingulate cortex. The IFG, sensitive to the SV/non-SV distinction, has been shown to increase activity in utilitarian in contrast to deontic judgments (Kahane et al., 2012), supporting the vision of SV/non-SV differences within the spectrum of utilitarian versus deontic judgment. Given the role of the IFG as prepotent response inhibition center (Aron, 2011), it is plausible that non-SV processing relies more heavily on cognitive control nodes facilitating a more utilitarian reasoning style over prepotent deontic judgments, in line with the increased reaction times observed in the non-SV condition. However, Berns et al. (2012) results, upon which we based our predictions, point to higher involvement of IFG activity in SVs compared to non-SVs, which the authors interpret as a higher reliance on rigid moral rule-retrieving when sacred values are invoked. Our MVPA data could support either of these results, since it reveals an information-rich cluster in the IFG for the classification between the SV and non-SV category. Nevertheless, the non-SV versus fixation cross results show an involvement of the IFG in non-SV processing, which is absent in the SV versus fixation cross contrast. In addition, our connectivity results reveal a down-regulation effect of the vmPFC/vACC cluster of activity on IFG activity in the sacred value condition. These observations lead us to think that, contrary to our predictions based on Berns' work (2012), activity in the IFG might be rather decreased during SV assessment compared to utilitarian/non-SV judgments, in line with Kahane's results.

Finally, activity in the precuneus was not initially predicted and thus, interpretation of this result is post-hoc. The precuneus is thought to regulate the transition between task-positive and task-negative states, connecting either with the right frontoparietal network during task-performance or with default mode network (DMN) nodes related to autobiographical memories during resting-state (Utevsky, Smith, & Huettel, 2014). According to the MVPA results, the precuneus was differentially involved in the non-SV and SV conditions, arguably switching between self-referential and task-positive states. In fact, whereas

typically executive frontoparietal circuits were more intensely recruited during non-SV trials (dlPFC, parietal cortex), SV trials involved rather task-negative regions, i.e. self-referential DMN nodes related to emotion engagement such as the vmPFC (Li, Mai, Liu, & Moran, 2014). Thus, it is plausible that SV assessment was more readily reliant on affective appraisal of autobiographical content in order to convey WFD responses, whereas non-SV judgments were rather sustained by task-positive frontoparietal circuitry, likely related to utilitarian analysis. Additionally, differential activity in the midcingulate cortex in SV versus non-SV classification, close to the posterior cingulate cortex, could be related to adopting the first-person perspective (Apps, Lockwood, Balsters, & Hunt, 2013).

Perceived spirituality of the out-group as predictor of frontoparietal activity

Among all participants, those who gave the US higher spiritual formidability ratings exhibited higher dlPFC/dmPFC and parietal activity while conveying WFD for both SVs and non-SVs. As mentioned above, these areas have been associated with executive attentional networks and utilitarian thinking and were more active during non-SV assessment. This finding suggests that participants who perceived higher out-group spirituality exhibited a neurofunctional profile closer to that of non-SV trials both during SV and non-SV trials. Why this was the case cannot be answered with the present data and, thus, further research is needed to elucidate this question. For now, one possible explanation could be that those who were able to acknowledge the out-group's spirituality were also the ones incurring in less out-group dehumanization. If dehumanization is defined as the act of attributing animalistic or machine-like traits to someone (Haslam, 2006), attributing uniquely human traits such as spiritual strength could be understood as form of humanization. Thus, it is plausible humanization of the out-group and, relatedly, lower levels of inter-group discrimination (Paladino, Rodriguez, Rodriguez, Gaunt, & Demoulin, 2002; Vaes, Leyens, Paladino, & Miranda, 2012), were associated with more deliberative responses to WFD for SVs and non-SVs. Similarly, higher dehumanization of the out-group (lower spirituality scores) would be associated with more straightforward deontic responses, and, relatedly, less frontoparietal recruitment in both SV and non-SV trials. Although this interpretation is highly speculative, the out-group physical formidability measure may serve as a partial control conveying some support to this

hypothesis. As a non-uniquely human trait, physical formidability was not expected to relate to out-group humanization and, accordingly, it did not have any predictive power over neural activity attributable to utilitarian or deontic reasoning.

5. 3. Connectivity findings in relation to WFD for sacred and non-sacred values

Critically, the vmPFC/vACC was found to down-regulate dlPFC and IFG activity during SV assessment. These areas exhibited decreased activation in SVs compared to non-SVs, pointing at the possibility that vmPFC/vACC was partly involved in such reduction. As explained above, the dlPFC and the IFG have been associated with deliberative thinking and behavioral inhibition, respectively, and their down-regulation might likely undermine the ability to engage in utilitarian reasoning, thus favoring deontic judgment. These results suggest that the vmPFC/vACC might not only contribute to boost the sense of personal control in the context of SV assessment, but also to attenuate ponderation mechanisms that otherwise enable cost-benefit analysis. Given that participants responded to willingness to fight and die for SVs, such boost in personal control should not be understood as a mechanism aligned with individual self-interest, but as a means to subserve the participant's cultural group, i.e. the moral community sharing values he considers more important than his own life.

The connectivity analyses revealed that vmPFC/vACC contributed to posterior cingulate cortex (PCC) activity, a default mode network region associated with episodic memory retrieval (Wheeler & Buckner, 2004). Although this result did not survive multiple comparisons correction, it supports previous results in moral decision-making, where the PCC has been associated with moral judgment (Moll, 2002) and deontic versus utilitarian judgment (Kahane et al., 2012). In addition, the PCC has been involved in 'egocentric' representations of visuospatial information (Burgess, Quayle, & Frith, 2001; Schmitz & Johnson, 2007; Vogeley et al., 2004) and evaluation of self-relevant information in conjunction with the vmPFC (Northoff et al., 2006; Schmitz & Johnson, 2006). Therefore, although it cannot be considered a main finding, the contribution of the vmPFC/vACC on PCC activity during SV trials can be interpreted as the adoption of an 'egocentric' perspective driven by retrieval of self-relevant episodic memories.

In the context of non-SVs, both dlPFC and dmPFC positively contributed to activity in the temporoparietal junction (TPJ), whereas the dlPFC also contributed to activity in the orbitofrontal cortex (OFC). Given the prominent role of the TPJ in mentalizing and perspective-taking, it is plausible that the TPJ, the dmPFC and the dlPFC work in conjunction to arrive to optimal responses after a process of back and forth perspective-switching and evaluation of behavioral outcomes. Interestingly, the vmPFC/vACC inhibited the dlPFC in SV trials, but we did not find the reverse pattern in non-SV trials, i.e. the dlPFC did not seem to inhibit affective responses associated with vmPFC/vACC activity (Li et al., 2014) in favor of utilitarian reasoning in non-SV trials. Therefore, it seems unlikely that the function of dlPFC in relationship to non-SVs is that of reappraisal of affective-laden stimuli, an otherwise common dlPFC function (Ochsner and Gross, 2005). Rather, it seems more likely that the more modest affective load of non-SVs might only poorly involve vmPFC/vACC activity, maybe due to an increased psychological distance to non-SVs. Thus, the apparent lack of prepotent affective responses at non-SVs might stimulate a longer more dispassionate search for appropriate responses driven by the dlPFC in front of researchers' queries. Moreover, contibution to OFC activity by dlPFC in non-SV trials might be related to the attribution of a motivational value to distinct behavioral options contrasted within the process of utilitarian reasoning (Kahane et al., 2012; Berns et al., 2012).

Experiment 2: Social influence on willingness to fight and die

In the Feedback paradigm, the main findings include: i) willingness to fight and die alignment with peers after feedback manipulation; ii) higher post-manipulation self-reported moral outrage at peers-lower feedback and higher joy at peers-agree feedback for SVs than for non-SVs; iii) parietal opercular activity was sensitive to the distinction between peers-lower versus peers-agree feedback for SVs and such responsiveness could be predicted by self-reported moral outrage in the peers-lower feedback for SV versus non-SV classification, and iv) individuals who changed their willingness to fight and die ratings for SVs to a greater extent also exhibited higher activity in cognitive control (IFG) and somatosensory regions (insula) while receiving peers-lower compared to peers-agree feedback for SVs.

5. 4. Behavioral results: reportedly angry but still susceptible to change

As predicted, the peers' feedback manipulation had a significant effect on participants WFD ratings, with participants changing WFD ratings in line with the group norm. Participants did not directly adopt community ratings; rather, they changed their judgment in the direction established by peers, similarly to Klucharev et al. (2009), who also used a 2 to 3 point difference between self and group ratings. Against expectations, sacred values did not show immunity to social influence compared to non-SVs. This is probably due to the fact that we did not directly measure changes in sacred value adherence. Changes in SV adherence could have been evaluated by checking if resistance to financial concessions and rejection of democratic consensus (the two SV criteria) hold true after a normative feedback manipulation on the same exact measures (financial concessions/democratic consensus). However, the present feedback manipulation and the correspondent change in judgment refer to willingness to fight and die (WFD) for SVs instead of SV adherence. The present results provide evidence that WFD is susceptible to social influence by the in-group, regardless of whether the value to defend is sacred or not. This fact allows for a greater range of possibilities with respect to potential de-radicalization interventions specifically targeting WFD rather than SV adherence, which has been observed to be resistant to social influence in other studies (Sheikh et al., 2013).

With respect to reported post-manipulation emotions towards the in-group, negative emotions (moral outrage and shame) were sensitive to peers-lower feedback, whereas positive emotions (compassion, joy and pride) increased in response to peers-agree feedback. All five emotions obtained significantly higher scores for SVs in contrast to non-SVs. These results are consistent with the notion that SVs engage emotional processes to a greater extent than non-SVs, reinforcing the distinction between emotion-driven deontic judgment and cognitively-driven utilitarian thinking. Moreover, negative and positive emotions where highly consistent among themselves and individuals showed consistent degrees of emotionality, with increased negative emotions towards peers-lower feedback correlated with enhanced positive emotions towards peers-agree feedback. Additionally, emotion ratings where somewhat related to WFD ratings, with WFD for SVs correlating with negative emotions at peers-lower feedback for non-SVs, and WFD for non-SVs as well as resistance to change WFD for SVs correlating with positive emotions at peers-agree feedback for non-SVs. Thus, expressed emotionality at peers feedback was

not only associated with value sacredness but also with WFD to some extent, pointing at the possible mediating role of emotion in the relationship between SV adherence and WFD.

5. 5. Neurofunctional findings on social influence over willingness to fight and die

According to the univariate analysis results, the only clusters responsive to peers-lower feedback for SVs compared to both peers-agree feedback for SVs and peers-lower feedback for non-SVs were located in the insula and the dlPFC. Although these results did not survive multiple comparisons correction ($p < .01$ uncorrected), they align with those of Klucharev et al (2009), who found dlPFC and insula activity in response to conflicting social opinions making use of a social feedback paradigm aimed to influence facial attractiveness ratings. The insula, especially in the anterior aspects, is thought to integrate somatosensory information from circumventing opercular cortices producing an emotional context embedding higher-order operations (Bechara et al., 2000). Specifically, the insula, together with the parietal operculum, have been involved with both physical and moral disgust responses (Calder, Keane, Manes, Antoun, & Young, 2000; Jabbi, Bastiaansen, & Keysers, 2008). Therefore, the present results are in line with the hypothesis that the insula might be responsive to conflicting social opinions, especially when the values at stake are sacred. Moreover, increased dlPFC activity at peers-lower feedback was found close to Brodmann area 9, a region that has been associated with increased self-criticism understood as higher error processing and behavioral inhibition (Longe et al., 2010).

Interestingly, change in WFD after peers-lower feedback for SV predicted activity in the anterior insula (AI), and the IFG. Again, these results did not survive multiple comparisons correction ($p < .001$ uncorrected). However, they partly replicate observations found by Berns et al. (2010), who detected AI activity in association with change in judgment to conform social input. AI activity could encode the costs of social exclusion by responding to socially aversive stimuli (Eisenberger, 2012). Additionally, predicted IFG activity could be related to increased behavioral control in participants who changed their WFD ratings the most. These results reinforce the notion of cognitive control areas working as mediators of change in judgment elicited by peers' influence. Hence, conflicting normative opinion might have sparked error-detection and behavioral

control processes to a greater extent in participants who were more sensitive to social feedback, highlighting the role of social information in reinforcement learning.

Interestingly, the fact that participants aligned to the social norm despite reporting feelings of moral outrage towards peers' ratings suggests that participants were probably unaware of their sensitivity to social feedback. Hence, alignment with peers seems like a rather automatic process escaping conscious decision-making. We did not measure neural activity during the second WFD rating and, thus, cannot confirm that the "public" behavioral shifts in WDF were also maintained within the personal domain in form of neurobiological changes (Zaki et al., 2011). However, incoherence between responsiveness to social input and moral outrage reports suggests that changes in WFD were not deliberate but responded to social expectations operating in a more subtle probably biologically-driven way. This interpretation is in line with Zaki et al. (2011) notion of an adaptive role of the social conformity bias in motivating consensus and coordinating action towards collective goals.

The multivariate pattern analysis provided more robust results, which survived to $p < .05$ FWE cluster-correction (single voxel $p < .001$). Specifically, differential activity in the left parietal operculum was found in the peers-lower compared to peers-agree condition for SVs. Given that both contrasted feedback conditions involved SVs, the left parietal operculum specifically responded to feedback type in the context of SVs. Interestingly, self-reported moral outrage at peers-lower feedback for SVs significantly predicted activity in the left parietal operculum while receiving peers-lower feedback for SVs versus non-SVs. Here, since feedback type was the same in both contrasted conditions, parietal operculum activity specifically responded to value type (SV/non-SV) in the context of peers-lower feedback as a function of self-reported moral outrage scores. However, no interaction effects were found between feedback and sacredness type.

The parietal operculum includes the secondary somatosensory cortex, which is densely connected with the insula and has been involved in perception of pain (Frot, Magnin, Mauguiere & Garcia-larrea, 2007; Bornhövd et al., 2002). In a study run by Sawamoto et al. (2000) the parietal operculum was related to both experience and expectation of thermal pain delivered by a CO2 laser stimulator. Peak MNI coordinates associated with both experienced pain and expected pain reported in that study are fairly close to the local

maxima observed in the present study. Other studies have found insula-opercular responsiveness to unpleasant odors (Wicker et al., 2003) and tastes (Rolls, 2015), highlighting the sensitivity of these areas to disgusting stimuli. In the social domain, the somatosensory operculum and the insula have been included in the so-called social aversion network (Bickart, Dickerson, & Feldman, 2014; Hayes & Northoff, 2011), which responds to social stimuli eliciting disgust, anger or indignation (Rilling, King-casas, & Sanfey, 2008; Sanfey et al., 2007) and is relevant for aversive conditioning in social reinforcement learning (Seymour et al., 2004). Moreover, Fukushima, Goto, Maeda, Kato, & Umeda (2013) found an association between judgment of agency (indistinctive between self and other-attributions) and the insula-opercular complex, which was interpreted as supporting bodily representations based on somatosensory information. Thereby, it is possible that witnessing conflicting in-group opinions triggered both judgments of agency dissociating one's opinion from the groups' and reactions of moral disgust and social avoidance, especially when values were sacred.

Finally, the bilateral parietal operculum has been associated with moral elevation (ME), i.e. feelings of self-transcendence induced by "witnessing acts of moral beauty" (Englander, Haidt, & Morris, 2012), in a paradigm consisting on videos with moral elevation content (e.g. "A man saves a stranger's life on the New York City subway") versus videos eliciting admiration (e.g. "Dancers compete on a talent competition"). Additionally, the bilateral parietal operculum was the only region exhibiting synchronized activity between participants watching the same ME clip as revealed by an inter-subject correlation method based on modelling BOLD activity using another participant's neural activity. At a speculative level, it could be that the observed opercular responsiveness to feedback type could be due to feelings of moral elevation at witnessing peers-agree in contrast to peers-lower feedback. No relationship was found between vlPFC activity and sensitivity to social feedback for SVs, contrary to Pincus et al. (2014) findings.

General discussion

This is the first neuroimaging study assessing the neural mechanisms behind willingness to fight and die for sacred values, and the first to assess sacred values in the context of real-world intergroup conflict. As shown by previous anthropological and socio-psychological research, sacred values can be understood as protected cultural views that work as core elements in intercultural conflict, confronting those who defend them with anyone who poses a threat to their integrity. As observed in the present study, people are ready to sacrifice themselves and even give up their lives in defense of such group-defining values, which can be both religious and nationalistic. According to several authors, sacred values attain a privileged position as they are adopted by moral community members as identitary elements, defining "who I am" primarily in terms of shared parochial sacred views. Such existential commitment to group-related sacred values is intensified by ritualistic activity, costly sacrifices and perceived existential threat by external forces, which increases feelings of brotherhood among members and group cohesiveness. As they turn into identity building blocks, sacred values may become more important to the beholder, who he is ready to fight and die to defend them. After all, if the beholder's identity is merged with that of his moral community, his "spirit" will survive as long as his community survives. Not surprisingly, when interactions with an opposing out-group threaten the integrity of sacred values, a spiral of tensions easily builds up, often crystallizing into an intractable conflict. Such deep-seated conflicts are frequently nourished and sustained by historical narratives, religious rhetoric, and an aura of in-group moral superiority.

In this sociocultural framework, the devoted actor model of decision-making aims to characterize individuals with a strong sense of group belongingness who react with moral outrage at violations of moral norms, rely on deontic judgment over utilitarian reasoning, and are ready to take action independently of risks and outcomes. The figure of the martyr would be the utmost expression of a devoted actor, who takes up costly sacrifices in favor of moral convictions related to the protection of a group of significant others. In this sense, the martyr is perceived in a positive almost divine light by his community. Portrayed as a hero, the martyr altruistically fights against the evil in an often oversimplified and highly abstract division of the good and the bad.

In the present study, we evaluated a set of predictions derived from the devoted actor model of decision-making, both at a behavioral and at a neurofunctional level:

1. Devoted actors will be willing to fight and die to defend their sacred values.
2. Devoted actors will diminish activity in brain regions associated with utilitarian reasoning in favor of unreflective deontic thinking when sacred values are at stake.
3. Devoted actors will be resistant to social pressure over willingness to fight and die for sacred values.
4. Devoted actors will respond with moral outrage to conflicting social input, which will be observable both at a behavioral and at a neural level.

As seen in the results section, all but prediction 3 was supported by our data. The present sample of Pakistani males with radicalized political views showed:

1. A fulminating expressed will to fight and die to defend sacred values such as "Prophet Mohammed must never be caricatured" or "US military should be expelled from Muslim lands" (mean of 6.59 out of 7) with very consistent ratings (sd = .49).
2. Reduced activity in brain areas associated with utilitarian thinking, deliberation and cognitive control (dlPFC, dmPFC and IFG), partly down-regulated by a strong activation of the vmPFC/vACC, associated with perceived control and self-agency.
3. No interaction between social feedback type and value sacredness was found, with social input influencing WFD for both sacred and non-sacred values.
4. Increased moral outrage at conflicting social opinions on WFD for sacred values, which predicted neural activity differences in the parietal operculum, associated with experienced and expected pain and moral disgust.

These results support the notion of a devoted actor who is willing to sacrifice for sacred values shared with his moral community, who largely relies on non-utilitarian deontic reasoning to solve trade-off scenarios pitching SVs against his own life, and who expresses rejection to conflicting social input, seemingly, as a means to protect SV integrity. Such observations are consistent with Greene et al. (2004) established affective-cognitive framework surrounding moral judgment presented in the introduction. According to this model, automatic affective processes to moral dilemmas respond to social expectations and are supported by self-referential activity in midline cortical structures (mPFC, PCC) and sociocognitve nodes (TPJ, STS). Such is the case of personal moral dilemmas, where people generally choose not to directly harm others even in

exchange for a greater utility, suggesting the use of a deontic rule-of-thumb (Greene et al., 2004). These prepotent socially-desirable responses compete with cognitive evaluative processes mediated by dlPFC, IFG and OFC, which are thought to dampen affective prepotent responses in favor of more distant goals involving greater utility.

According to this interpretation, fighting and dying for SVs can be understood as a socially-modulated affective-driven outcome sustained by activity in ventromedial portions of the CMS (vmPFC/vACC) and TPJ. Conversely, WFD for non-SVs activated areas that parallel Greene's observed neural response to impersonal moral dilemmas (dlPFC, IFG, OFC), pointing at a higher involvement of cognitive evaluation processes. Thus, whereas SV assessment seems to rely on deontic thinking responding to social expectations, non-SV assessment calls upon rather utilitarian modes of decision-making. Again, these interpretations depict the devoted decision-maker as a figure who is deeply moved by social expectation and applies learned deontic rules to solve upcoming moral dilemmas invoking SVs. Altoghether, WFD scores for SVs are processed by neural nodes, i.e. vmPFC/ACC, related to self-initiated action (Haynes et al., 2015) and perceived control (Hashimoto et al., 2015) that, in turn, are responsive to social pressure (Izuma et al., 2008), integrating socially desirable outcomes into personal goal-oriented drive (Cunningham & Zelazo, 2007). In the light of this evidence, it is not surprising that WFD for SVs turned out to be sensitive to social feedback, pointing to a possible relationship between increased sense of power through social validation by the moral community.

Despite that Greene et al. (2004) morality model accommodates most of our findings, it fails to establish a neurofunctional difference between dorsal and ventral portions of the medial prefrontal cortex, a well-established distinction in the self-referentiality literature that seems to be at the core of the SVs/non-SV dichotomy. On one side, the vmPFC-ACC-subcortical network has been pointed out as the 'affective' pathway of self-referentiality (Schmitz & Johnson, 2006) generating a pre-attentive bias towards explicit self-relevant information based on anticipation and mnemonic content (Schmitz & Johnson, 2007). Along similar lines, Andrews-Hanna, Reidler, Sepulcre, Poulin, & Buckner (2010) identify the vmPFC-middle temporal system as a subdivision of the default mode network (DMN). The dense connections between vmPFC and emotion processing nodes (insula and amygdala) are thought to provide an affective context in social cognition settings, using mnemonic resources to construct "mental scenes". In turn, the generated mental

scenes might have an adaptive role in improving prediction accuracy based on current affective states (Andrews-Hanna et al., 2010). Therefore, the vmPFC/vACC together with middle temporal activity observed in SV trials could be well explained by the elicitation of mental scenes generating an affective context bound to the self.

On the other side, the dmPFC-cortical-subcortical network has been proposed as the 'cognitive' pathway of self-referentiality (Schmitz and Johnsson, 2006), which is thought to sustain introspective processes by recollecting, manipulating and evaluating self-relevant content (Schmitz, 2007). In parallel, Andrews-Hanna et al. (2010) point at the dmPFC subdivision of the DMN as a functionally distinct self-referential network activated during self-relevant decisions that demand mental state inferences (Lieberman, 2007; Olsson & Ochsner, 2007) and metacognitive processes (Ochsner et al., 2004). Thus, increased dmPFC recruitment in non-SV trials might be triggering introspective and evaluative processes, which are, in turn, silenced during SV trials. In contrast to the reproduction of rigid internalized social norms, such introspective processes might allow a more flexible response to changing contexts, possibly facilitating the adoption of more utilitarian outcome-oriented responses. Despite of the potential role of self-evaluation and introspection mediated by dmPFC in moral judgment, this aspect of social cogntion is not clearly delimited in Greene's et al. (2004) model.

In sum, willingness to fight and die for SVs involved brain regions identified in both Greene's moral judgment model (2004) as well as in self-referential/DMN models of cognition. Concretely, WFD for SVs (in contrast to non-SVs) seems to elicit 1) vmPFC/vACC-processed affective responses biased by socially learned "rights" and "wrongs", possibly stored and retrieved in form of emotion-laden mental scenes, thereby articulating deontic thinking in a cognitively rigid way, 2) the down-regulation of introspective/metacognitive processing centers (dmPFC) and executive frontoparietal nodes supporting flexible utility-maximizing decision-making (dlPFC, parietal cortex), partly due to vmPFC/vACC activity. In other words, whereas WFD for SVs seems to be dominated by affective self-referential activity (vmPFC subdivision of the DMN), WFD for non-SVs seems to involve a mix of neural activity associated with cognitive self-referentiality (dmPFC subdivision of the DMN) and executive frontoparietal functions (dlPFC, parietal cortex).

Finally, it is interesting to note that the motives that seem to characterize devoted actor decision-making can be paradoxically seen both as altruistic and egotistical, depending on the social context where they are evaluated. As previously known and confirmed here, the devoted actor is ready to sacrifice his life in order to preserve his people's views in an act of extreme altruism. However, the present results suggest that, in doing so (or imagining doing so) he experiences a deep sense of control and self-agency that motivates his actions at the expense of deliberation, independently of risks and outcomes. In turn, if the act of fighting and dying damages or violates competing SVs (e.g. lives of innocent civilians), the devoted actor imposes an egocentric bias by forcing his group's views as absolute truths. Therefore, what can be seen as a selfless altruistic act can, in turn, function as an empowerment tool that, despite of being poorly synchronized with the martyr's physical well-being, is highly biased towards the self and his moral community. It is important to note here that moral community is employed to designate the immediate social environment sharing values and beliefs rather than extended religious communities.

Three possible pathways to de-radicalization emerge directly or indirectly from the present results. In the first place, the increased dmPFC/dlPFC activity shown in non-SV trials associated with lower WFD scores supports de-radicalization strategies based on cultivating cognitive flexibility, critical thinking and perspective-taking, which already count on some empirical support (Hameiri et al., 2014; Galinsky & Moskowitz, 2000). Moreover, and as a direct implication of this study, the decrease in WFD by means of the social feedback manipulation suggests that de-radicalization strategies could benefit from targeting willingness to fight and die rather than SV adherence. Given that means to achieve goals (fighting and dying) seem more sensitive to external influences than goals *per se* (Sheikh et al., 2013), providing credible alternative approaches to pursue personal goals could be a reasonable starting point. In relationship to this, and given the reliance of WFD on social expectations and learned deontic content (what's "right" and "wrong"), its seems crucial to socially reinforce such alternative means of action. However, for this approach to work out, individuals need to *care* about the opinion of the social group reinforcing their personal goals in the first place. For that, it seems a requirement that they feel integrated and supported within more extended social networks co-inhabiting within the established sociopolitical system. Providing individuals a sense of social responsibility, allowing them to make relevant life choices in a scenario of real opportunities and facilitating tools for real political participation might help reduce the

need to engage in what can be seen as extreme acts of social validation and self-empowerment.

5. 6. Limitations

Given the difficulties in recruiting a sample of Pakistani males with radicalized political views in the area of Barcelona, the sample size was limited, making the analysis statistically more challenging. From the perspective of the experimental paradigm, several limitations could be identified. On one side, due to the fact that social feedback was calculated based on previous participants' WFD ratings and given that most of SV ratings were at ceiling, the peers-higher feedback condition could only be rarely implemented and thus had to be dropped from the analysis. Moreover, SVs and non-SVs were presented together with the response Likert scale. This decision had the advantage to keep participants attentive, but also added some noise to the BOLD signal associated with SVs and non-SVs, especially preparatory motor activity to move the cursor to a previously known direction, eye gaze movement along the scale and the visual content associated with the scale. Indeed, participants seemed to be attentive during the task, as shown by the above chance classification accuracy between different randomly chosen values based on left middle temporal activity, an area associated with recognition of semantic content while reading (Acheson & Hagoort, 2013).

The presence of some non-responded trials forced us to work with a slightly unbalanced data set, which can bias neuroimaging first-level analysis results towards the most frequent condition for each participant. Other BOLD signal confounds in the MVPA analysis could be identified using the SAA method that could successfully classify SVs and non-SVs based on participants' WFD ratings, and to a lesser extent, reaction times. This makes it difficult to distinguish neural activity associated independently with SVs and WFD. Thus, our results refer specifically to WFD for SVs, but not SVs or WFD alone. In addition, reaction time differences between SV and non-SV responses constitute a modest but significant confound. Hence, we cannot distinguish neural substrates of WFD for non-SVs and longer reaction times, which we attribute to deliberation processes intrinsic to the response to WFD for non-SVs.

5. 7. Conclusions

This is the first neuroimaging study targeting willingness to fight and die for sacred values in a real-world sample within the context of intergroup conflict. Our results suggest that WFD to defend sacred values is prominently sustained by vmPFC/vACC activity, associated with self-agency and affective appraisal of self-referential content, thus awakening a sense of power and control. These processes run at the expense of deliberative thinking and perspective-taking activity, partly through the down-regulation of dlPFC and IFG by vmPFC/vACC. Thus, self-positioning is stable and immediate, responding to socially-acquired deontic judgments of "right" and "wrong". In contrast, decisions involving non-sacred values seem to be more subject to ponderation and cost-benefit analysis mediated by dlPFC, dmPFC, IFG and TPJ activity, calling upon perspective-taking processes in order to find suitable responses. Furthermore, social influence appears as an important moderating factor of WFD. Even though conflicting social opinions provoked both emotional reactions of moral outrage and neural activity associated with pain and disgust in the parietal operculum, participants still exhibited a significant change in WFD in the direction established by the social norm. The observed moral outrage responses to conflicting social influences together with positive emotions such as joy at agreeing social input might have a reinforcing effect on one's own beliefs, ultimately promoting, protecting and securing their integrity. Therefore, attachment to sacred values might trigger an egocentric bias, where a set of selected SVs are imposed over competing SVs by means of fighting and dying within acts of blind self-empowerment. Our results suggest that, rather than dismantling SV adherence, de-radicalization strategies could benefit from targeting WFD, thus emphasizing and socially reinforcing other means to achieve personal goals. In addition, for social reinforcement to take place within the context of co-existing moral communities it seems crucial that individuals feel socially included within larger social networks supported by the established sociopolitical system, thus reinforcing a sense of responsibility and care for common ground values.

References

Acheson, D. J., & Hagoort, P. (2013). Stimulating the brain's language network: syntactic ambiguity resolution after TMS to the inferior frontal gyrus and middle temporal gyrus. *J Cogn Neurosci.* 2013 Oct; 25(10):1664-77. http://doi.org/10.1162/jocn

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Publishing Group*, *14*(10), 1338–1344. http://doi.org/10.1038/nn.2921

Alexander, L. and Moore, M., "Deontological Ethics", *The Stanford Encyclopedia of Philosophy* (Spring 2015 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2015/entries/ethics-deontological/>.

Allison, T., Puce, A., & Mccarthy, G. (2000). Social perception from visual cues: role of the STS region, *Trends on Cognitive Sciences*, 4(7).

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds : the medial frontal cortex and social cognition, Nature Reviews, *7*, 268–277. http://doi.org/10.1038/nrn1884

Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., & Buckner, R. L. (2010). Functional-anatomic fractionation of the brain's default network. *Neuron.* 2010 Feb 25;65(4):550-62. doi: 10.1016/j.neuron.2010.02.005.

Apps, M. A. J., Lockwood, P. L., Balsters, J. H., & Hunt, L. T. (2013). The role of the midcingulate cortex in monitoring others ' decisions, *Front Neurosci.* 2013 Dec 20;7:251. doi: 10.3389/fnins.2013.00251.

Argembeau, A. D. (2013). On the role of the ventromedial prefrontal cortex in self-processing : the valuation hypothesis, *Frontiers in Human Neuroscience*, *7*, 372. http://doi.org/10.3389/fnhum.2013.00372

Aron, A. R. (2011). From Reactive to Proactive and Selective Control: Developing a Richer Model for Stopping Inappropriate Responses. *Biological Psychiatry*, *69*(12), e55–e68. http://doi.org/10.1016/j.biopsych.2010.07.024

Asch, S. E. (1955). Opinions And Social Pressure, *Scientific American* November 1955 Vol. 193, No. 5, pp. 31-35.

Atran, S. (2003). Genesis of suicide terrorism. *Science (New York, N.Y.)*, *299*(5612), 1534–1539. http://doi.org/10.1126/science.1078854

Atran, S. (2010). The Evolution of Religion: How Cognitive By-Products, Adaptive Learning Heuristics, Ritual Displays , and Group Competition Generate Deep Commitments to Prosocial Religions, *Biological Theory* March 2010, *5*(1), 18–30.

Atran, S. (2006) Devoted Actor Versus Rational Actor Models for Understanding World Conflict," *The National Security Council*, 14 Sep. 2006.

Atran, S., & Axelrod, R. (2008). In Theory: Reframing Sacred Values, *Negotiation Journal* (July), 221–246.

Atran, S., Axelrod, R., & Davis, R. (2007). Sacred barriers to conflict resolution. *Science (New York, N.Y.)*, *317*(5841), 1039–1040. http://doi.org/10.1126/science.1144241

Atran, S., & Ginges, J. (2012). Religious and Sacred Imperatives in Human Conflict. *Science*, *336*(6083), 855–857. http://doi.org/10.1126/science.1216902

Barlow, D. H., Ellard, K. K., Sauer-Zavala, S., Bullis, J. R., & Carl, J. R. (2014). The Origins of Neuroticism. *Perspectives on Psychological Science* September 2014 vol. 9 no. 5 481-496. http://doi.org/10.1177/1745691614544528

Baron, J., & Leshner, S. (2000). How serious are expressions of protected values? *Journal of Experimental Psychology: Applied*, *6*(3), 183–194. http://doi.org/10.1037//1076-898X.6.3.183

Baron, J., & Ritov, I. (2009). Chapter 4 Protected Values and Omission Bias as Deontological Judgments. *Psychology of Learning and Motivation*. http://doi.org/10.1016/S0079-7421(08)00404-0

Baron, J., & Spranca, M. (1997). Protected Values. *Organizational Behavior and Human Decision Processes*, *70*(1), 1–16. http://doi.org/10.1006/obhd.1997.2690

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion , Decision Making and the Orbitofrontal Cortex, *Cereb Cortex.* 2000 Mar;10(3):295-307.

Bechara, A., Tranel, D., Damasio, H., & Damasio, R. (1996). Failure to Respond Autonomically to Anticipated Future Outcomes Following Damage to Prefrontal Cortex, *Cereb Cortex.* 1996 Mar-Apr;6(2):215-25.

Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Fs, M. (2009). Associative learning of social value, *Nature* 456, 245-249. doi:10.1038/nature07538

Bering. J.M., & Johnson, D.D.P. (2005). "O Lord . . . You Perceive my Thoughts from Afar": Recursiveness and the Evolution of Supernatural Agency∗,*Journal of Cognition and Culture* 5.1-2.

Bernhardt, B. C., & Singer, T. (2012). The Neural Basis of Empathy. *Annu Rev Neurosci.* 2012;35:1-23. doi: 10.1146/annurev-neuro-062111-150536.

Berns, G. S., Bell, E., Capra, C. M., Prietula, M. J., Moore, S., Anderson, B., … Atran, S. (2012). The price of your soul: neural evidence for the non-utilitarian representation of sacred values. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1589), 754–762. http://doi.org/10.1098/rstb.2011.0262

Berns, G. S., Chappelow, J., Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., & Richards, J. (2005). Neurobiological Correlates of Social Conformity and Independence During Mental Rotation. *Biological Psychiatry*, *58*(3), 245–253. http://doi.org/10.1016/j.biopsych.2005.04.012

Bickart, K. C., Dickerson, B. C., & Feldman, L. (2014). Neuropsychologia The amygdala as a hub in brain networks that support social life, *Neuropsychologia.* 2014 Oct;63:235-48. doi: 10.1016/j.neuropsychologia.2014.08.013.

Blair, R. J. R. (2004). The roles of orbital frontal cortex in the modulation of antisocial behavior, *Brain Cogn.* 2004 Jun;55(1):198-208. http://doi.org/10.1016/S0278-2626(03)00276-8

Blair, R.J.R (2001). Neurocognitive models of aggression, the antisocial personality disorders , and psychopathy, *J Neurol Neurosurg Psychiatry* 2001;71:727–731.

Bollini, A. M., Walker, E. F., Hamann, S., & Kestler, L. (2004). The influence of perceived control and locus of control on the cortisol and subjective responses to stress responses to stress, *Biol Psychol*. 2004 Nov;67(3):245-60.

Borg, J. S., Hynes, C., Horn, J. Van, & Grafton, S. (2006). Consequences, Action, and Intention as Factors in Moral Judgments: An fMRI Investigation, *J Cogn Neurosci.* 2006 May;18(5):803-17. http://doi.org/10.1162/jocn.2006.18.5.803.

Bornhövd, K., Quante, M., Glauche, V., Bromm, B., Weiller, C., Büchel, C. (2002). Painful stimuli evoke different stimulus ± response functions in the amygdala , prefrontal , insula and somatosensory cortex : a single-trial fMRI study, *Brain.* 2002 Jun;125(Pt 6):1326-36.

Boyd, R., & Richerson, P.J. (1985). Culture and the Evolutionary Process. *University of Chicago Press* (Chicago).

Brass, M., & Haggard, P. (2007). To Do or Not to Do: The Neural Signature of Self-Control, J *Neurosci.* 2007 Aug 22;27(34):9141-5.

Brass, M., & Haggard, P. (2016). The What , When , Whether Model of Intentional Action, *Neuroscientist.* 2008 Aug;14(4):319-25. doi: 10.1177/1073858408317417.

Broome, J. (1991). UTILITY, *Economics and Philosophy*, 7 (1):1-12.

Brothers L, Ring B. (1992) A neuroethological framework for the representation of minds. *J Cogn Neurosci*. *1992 Spring*;*4*(*2*):*107-18. doi*: *10.1162*/*jocn.1992.4.2.107.*

Brower, M. C., Price, B. H., & Program, S. C. (2001). Neuropsychiatry of frontal lobe dysfunction in violent and criminal behaviour : a critical review, *J Neurol Neurosurg Psychiatry* 2001;71:720–726

Buckholtz, J. W., Callicott, J. H., Kolachana, B., Hariri, A. R., Goldberg, T. E., Genderson, M., & Egan, M. F. (2008). Genetic variation in MAOA modulates ventromedial prefrontal circuitry mediating individual differences in human personality, *Molecular Psychiatry* (2008) 13, 313–324; doi:10.1038/sj.mp.4002020

Buckner, R. L., Andrews-Hannah, J., & Schacter, D.L. (2008). The Brain ' s Default Network Anatomy, Function, and Relevance to Disease, *Ann. N.Y. Acad. Sci.* 1124: 1–38. http://doi.org/10.1196/annals.1440.011

Burgess, P. W., Quayle, A., & Frith, C. D. (2001). Brain regions involved in prospective memory as determined by positron emission tomography, *Neuropsychologia*, 2001;39(6):545-55.

Burgess, H. & Burgess, G. "What Are Intractable Conflicts?" Beyond Intractability. Ed. Guy Burgess and Heidi Burgess. Conflict Research Consortium, University of Colorado, Boulder, Colorado, USA. Nov 2003. <http://www.beyondintractability.org/essay/meaning-intractability>.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex, *Trends Cogn Sci.* 2000 Jun;4(6):215-222.

Calder, A. J., Keane, J., Manes, F., Antoun, N., & Young, A. W. (2000). Impaired recognition and experience of disgust following brain injury, *Nature Neuroscience* 3, 1077 - 1078 (2000)

Calder, A. J., Lawrence, A. D., & Young, A. W. (2001). Neuropsychology Of Fear And Loathing, *Nat Rev Neurosci.* 2001 May;2(5):352-63.

Calhoun LG, Cheney T, Dawes AS (1974) Locus of Control, Self-Reported Depression, and Perceived Causes of Depression. *Journal of Consulting and Clinical Psychology* 42(5):736 · November 1974

Campbell-meiklejohn, D. K., Bach, D. R., Roepstorff, A., Dolan, R. J., & Frith, C. D. (2010). Report How the Opinion of Others Affects Our Valuation of Objects. *Current Biology*, *20*(13), 1165–1170. http://doi.org/10.1016/j.cub.2010.04.055

Campbell-meiklejohn, D. K., Simonsen, A., Jensen, M., Wohlert, V., Gjerl, T., Scheel-kruger, J., … Roepstorff, A. (2012). Modulation of Social Influence by Methylphenidate, *Neuropsychopharmacology*. 2012 May;37(6):1517-25. doi: 10.1038/npp.2011.337.

Carrington, S. J., Bailey, A. J., Carrington, S. J., & Bailey, A. J. (2009). Carrington SJ , Bailey AJ . Are there theory of mind regions in the brain ? A review of the neuroimaging literature . *Hum Brain Mapp* 30: Are There Theory of Mind Regions in the Brain ? A Review of the Neuroimaging Literature, (July 2016), 2313–2335. http://doi.org/10.1002/hbm.20671

Chein, J., Albert, D., Brien, L. O., Uckert, K., & Steinberg, L. (2012). Peers increase adolescent risk taking by enhancing activity in the brain's reward circuitry, *Dev Sci.* 2011 March ; 14(2): F1–F10. doi:10.1111/j.1467-7687.2010.01035

Cheng, C., Chio, J. H., & Chan, M. S. (2012). Cultural Meaning of Perceived Control : A Meta-Analysis of Locus of Control and Psychological Symptoms Across 18 Cultural Regions. *Psychol Bull.* 2013 Jan;139(1):152-88. doi: 10.1037/a0028596.

Chun, M. M., Golomb, J. D., & Turk-browne, N. B. (2011). A Taxonomy of External and Internal Attention. *Annu Rev Psychol.* 2011;62:73-101. doi: 10.1146/annurev.psych.093008.100427.

Cialdini, R. B., & Goldstein, N. J. (2004). Social Influence : Compliance and Conformity, *Annu Rev Psychol.* 2004;55:591-621. http://doi.org/10.1146/annurev.psych.55.090902.142015

Clément, F. (2010). To Trust or not to Trust ? Children' s Social Epistemology. http://doi.org/10.1007/s13164-010-0022-3

Cloninger C, Svrakic D, Przybeck T (1993) A psychobiological model of temperament and character. *Arch Gen Psychiatry 50*: 975–990.

Cohen, M. X., & Ranganath, C. (2007). Reinforcement Learning Signals Predict Future Decisions, *J Neurosci.* 2007 Jan 10;27(2):371-8.

Corriveau, K. H., Fusaro, M., & Harris, P. L. (2009). Going With the Flow Preschoolers Prefer Nondissenters as Informants, *Association for Psychological Science*, 20(3), 372–377.

Cox, R.W., (1996) AFNI: software for analysis and visualization of func- tional magnetic resonance neuroimages, *Comput. Biomed. Res.* 29 (1996) 162–173.

Critchley, H. D., Mathias, C. J., Josephs, O., Doherty, J. O., Zanini, S., Dewar, B., … Dolan, R. J. (2003). Human cingulate cortex and autonomic control : converging neuroimaging and clinical evidence. *Brain* (2003),126, 2139±2152 http://doi.org/10.1093/brain/awg216

Croxson, P. L., Walton, M. E., Reilly, J. X. O., Behrens, T. E. J., & Rushworth, M. F. S. (2009). Effort-Based Cost – Benefit Valuation and the Human Brain, *The Journal of Neuroscience*, 8 April 2009, 29(14): 4531-4541; *doi: 10.1523/JNEUROSCI.4515-08.2009*

Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations : a social cognitive neuroscience perspective, *11*(3). *Trends Cogn Sci.* 2007 Mar;11(3):97-104 http://doi.org/10.1016/j.tics.2006.12.005

Darwin, C. (1874). The descent of man: And selection in relation to sex. (Revised edition.). Philadelphia: J.Wanamaker.

Decety, J., & Lamm, C. (2016). The Role of the Right Temporoparietal Junction in Social Interaction : How Low-Level Computational Processes Contribute to Meta-Cognition, *Neuroscientist.* 2007 Dec;13(6):580-93. http://doi.org/10.1177/1073858407304654

Dehghani, M., Atran, S., Iliev, R., Sachdeva, S., Medin, D., & Ginges, J. (2010). Sacred values and conflict over Iran ' s nuclear program, *Judgment and Decision Making*, Vol. 5, No. 7, December 2010, pp. 540–546

Dehghani, M., Iliev, R., Atran, S., Ginges, J., & Medin, D. (2009). Emerging sacred values: The Iranian nuclear program. *Judgment and Decision Making*, *4*(7), 990–993.

Deutsch M & Gerard H B. (1955) A study of normative and informational social influences upon individual judgment. *J. Abnorm. Soc. Psychol.* 51:629-36, 1955.

Dixon, M. L., & Christoff, K. (2014). Neuroscience and Biobehavioral Reviews The lateral prefrontal cortex and complex value-based learning and decision making. *Neuroscience and Biobehavioral Reviews*, *45*, 9–18. http://doi.org/10.1016/j.neubiorev.2014.04.011

Doherty, J. P. O. (2004). Reward representations and reward-related learning in the human brain : insights from neuroimaging, *Curr Opin Neurobiol.* 2004 Dec;14(6):769-76. http://doi.org/10.1016/j.conb.2004.10.016

Dreher, J., Kohn, P., Kolachana, B., Weinberger, D. R., & Faith, K. (2009). Variation in dopamine genes influences responsivity of the human reward system, *Proc Natl Acad Sci U S A.* 2009 Jan 13;106(2):617-22. doi: 10.1073/pnas.0805517106.

Duc, C., Hanselmann, M., Boesiger, P., & Tanner, C. (2013). Sacred values: Trade-off type matters. *Journal of Neuroscience, Psychology, and Economics*, *6*(4), 252–263. http://doi.org/10.1037/npe0000014

Durkheim, E. (1912) Elementary forms of Religious Life. *Free Press*, New York.

Eisenberger, N. I. (2012). The pain of social disconnection : examining the shared neural underpinnings of physical and social pain, *Nat Rev Neurosci.* 2012 May 3;13(6):421-34. doi: 10.1038/nrn3231.

Eisenberger, N. I., Way, B. M., Taylor, S. E., Welch, W. T., & Lieberman, M. D. (2007). From the Brain ' s Response to Social Exclusion, *Biol Psychiatry* 2007;61:1100–1108. http://doi.org/10.1016/j.biopsych.2006.08.007

Ellis, B. J., & Boyce, W. T. (2008). Biological Sensitivity to Context, *Dev Psychopathol.* 2005 Spring;17(2):271-301.

Englander, A., Haidt, J., & Morris, J. P. (2012). Neural Basis of Moral Elevation Demonstrated through Inter-Subject Synchronization of Cortical Activity during Free-viewing, *PLoS ONE* 7(6): e39384. doi:10.1371/journal.pone.0039384

Etkin, A., Egner, T., & Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in Cognitive Sciences*, *15*(2), 85–93. http://doi.org/10.1016/j.tics.2010.11.004

Falk, E. B., Way, B. M., & Jasinska, A. J. (2012). An imaging genetics approach to understanding social influence. *Frontiers in Human Neuroscience* (June), 1–13. http://doi.org/10.3389/fnhum.2012.00168

Farrow, T. F. D., Zheng, C. A. Y., Wilkinson, I. D., … Woodruff, P. W. R. (2001) Investigating the functional Investigating the functional anatomy of empathy and forgiveness, *Neuroreport.* 2001 Aug 8;12(11):2433-8. http://doi.org/10.1097/00001756-200108080-00029

Fast, N.J., Gruenfeld, D.H., Sivanathan, N., Galinsky, A.D. (2009). Illusory Control: A Generative Force Behind Power's Far-Reaching Effects. Stanford University Graduate School of Business Research Paper No. 2009.

Fischer, R. (2013). How Do Rituals Affect Cooperation ?, *Hum Nat.* 2013 Jun;24(2):115-25. doi: 10.1007/s12110-013-9167-y.

Fiske, A. P., & Tetlock, P. E. (1997). Taboo Trade-offs : Reactions to Transactions That Transgress the Spheres of Justice, *Political Psychology*, 18(2), 255–298.

Frith, C. D., & Frith, U. (1999). Interacting Minds — A Biological Basis, *Science.* 1999 Nov 26;286(5445):1692-5.

Frot, M., Magnin, M., Garcia-larrea, L., & F-, B. (2007). Human SII and Posterior Insula Differently Encode Thermal Laser Stimuli, *Cereb Cortex.* 2007 Mar;17(3):610-20. http://doi.org/10.1093/cercor/bhk007

Fukushima, H., Goto, Y., Maeda, T., Kato, M., & Umeda, S. (2013). Neural Substrates for Judgment of Self-Agency in Ambiguous Situations, PLoS ONE 8(8): e72267. doi:10.1371/journal.pone.0072267.

Gayer, C. C., Halperin, E., & Bar-tal, D. (2009). Overcoming Psychological Barriers to Peaceful Conflict Resolution The Role of Arguments about Losses, *Journal of Conflict Resolution* December 2009 vol. 53 no. 6 951-975

Gervais, W. M., & Norenzayan, A. (2012). Analytic Thinking Promotes Religious Disbelief, *Science* 27 Apr 2012: Vol. 336, Issue 6080, pp. 493-496 DOI: 10.1126/science.1215647

Ginges, J., Atran, S. (2010). What Motivates Participation in Violent Political Action : Selective Incentives or Parochial Altruism ? *Annals of the New York Academy of Sciences*, Wiley, 2009, 1167, pp.115-123. http://doi.org/10.1111/j.1749-6632.2009.04543.x

Ginges, J., Atran, S., Medin, D., & Shikaki, K. (2007). Sacred bounds on rational resolution of violent political conflict. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(18), 7357–7360. http://doi.org/10.1073/pnas.0701768104

Ginges, J., Hansen, I., & Norenzayan, A. (2009). Religion and Support for Suicide Attacks, *Psychol Sci.* 2009 Feb;20(2):224-30. doi: 10.1111/j.1467-9280.2009.02270.x

Gordon Marshall. "Moral Community." A Dictionary Of Sociology. 1998. Retrieved June 27, 2016 From Encyclopedia.Com: Http://Www.Encyclopedia.Com/Doc/1o88-Moralcommunity.Html

Görgen, K., Hebart, M. N., Allefeld, C., & Haynes, J.-D. (in preparation). Detecting, Avoiding and Eliminating Confounds in Neuroimaging Data Analysis: Design–Analysis Interactions and the Same Analysis Approach.

Grafman J, Schwab K, Warden D, Pridgen A, Brown HR, Salazar AM.(1996) Frontal lobe injuries, violence, and aggression: a report of the Vietnam Head Injury Study. Neurology. May;46(5):1231-8.

Greene, J.D. (2015). The Cognitive Neuroscience of Moral Judgment and Decision Making, *MIT Press* 1013–1024.

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, D. (2009). Cognitive Load Selectively Interferes with Utilitarian Moral Judgment. *Cognition*. 2008 June ; 107(3): 1144–1154.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The Neural Bases of Cognitive Conflict and Control in Moral Judgment, *Neuron*, *44*, 389–400.

Greene, J. D., Sommerville, R. B., & Nystrom, L. E. (2001). An fMRI Investigation of Emotional Engagement in Moral Judgment, *Science, 293*(September), 2105–2109.

Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work ?, *Trends Cogn Sci.* 2002 Dec 1 *6*(12), 517–523.

Haber, S. N., & Knutson, B. (2009). The Reward Circuit : Linking Primate Anatomy and Human Imaging. *Neuropsychopharmacology*, *35*(1), 4–26. http://doi.org/10.1038/npp.2009.129

Hakamata, Y., Iwase, M., Kato, T., Senda, K., & Inada, T. (2013). The Neural Correlates of Mindful Awareness : A Possible Buffering Effect on Anxiety-Related Reduction in Subgenual Anterior Cingulate Cortex Activity, PLoS ONE 8(10): e75526. doi:10.1371/journal.pone.0075526

Hameiri, B., Porat, R., Bar-tal, D., Bieler, A., & Halperin, E. (2014). Paradoxical thinking as a new avenue of intervention to promote peace, *Proc Natl Acad Sci U S A, 111*(30). http://doi.org/10.1073/pnas.1407055111

Hanselmann, M., & Tanner, C. (2008). Taboos and conflicts in decision making: Sacred values, decision difficulty, and emotions. *Judgment and Decision Making*, *3*(1), 51–63. Retrieved from http://www.sjdm.org/~baron/journal/bb5.pdf

Harnett, N. G., Wheelock, M. D., Wood, K. H., Ladnier, J. C., Mrug, S., & Knight, D. C. (2015). NeuroImage Affective state and locus of control modulate the neural response to threat. *NeuroImage*, *121*, 217–226. http://doi.org/10.1016/j.neuroimage.2015.07.034

Hashimoto, T., Takeuchi, H., Taki, Y., Sekiguchi, A., Nouchi, R., & Kotozaki, Y. (2015). NeuroImage Neuroanatomical correlates of the sense of control : Gray and white matter volumes associated with an internal locus of control. *NeuroImage*, *119*, 146–151. http://doi.org/10.1016/j.neuroimage.2015.06.061

Haslam, N., & Haslam, N. (2006). Dehumanization : An Integrative Review, *Pers Soc Psychol Rev.* 2006;10(3):252-64. http://doi.org/10.1207/s15327957pspr1003

Hayes, D. J., & Northoff, G. (2011). Identifying a network of brain regions involved in aversion-related processing : a cross-species translational investigation, *Frontiers in Integrative Science*, *5*(October), 1–21. http://doi.org/10.3389/fnint.2011.00049

Haynes, J.D., Wisniewski, D., Görgen, K., Momennejad, I., Reverberi, C. (2015). FMRI decoding of intentions: Compositionality , hierarchy and prospective memory, Conference: 3rd International Winter Conference on Brain-Computer Interface (BCI) http://doi.org/10.1109/IWW-BCI.2015.7073031

Hebart, M.N., Görgen, K. and Haynes, J.D. (2015). The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. *Frontiers in neuroinformatics* (January), 1–18. http://doi.org/10.3389/fninf.2014.00088

Henrich, J. (2009). The evolution of costly displays , cooperation and religion : credibility enhancing displays and their implications for cultural evolution. *Evolution and Human Behavior*, *30*(4), 244–260. http://doi.org/10.1016/j.evolhumbehav.2009.03.005

Hogeveen, J., Obhi, S. S., Banissy, M. J., Santiesteban, I., Press, C., Catmur, C., & Bird, G. (2015). Task-dependent and distinct roles of the temporoparietal junction and inferior frontal cortex in the control of imitation. *Soc Cogn Affect Neurosci.* 2015 Jul;10(7):1003-9. doi: 10.1093/scan/nsu148.

Hsu, M., & Quartz, S. R. (2008). The Right and the Good : Distributive ustice and neural encoding of equity and efficiency, *Science.* 2008 May 23;320(5879):1092-5. doi: 10.1126/science.1153651.

Hutcherson, C. A., Plassmann, H., Gross, J. J., & Rangel, A. (2012). Cognitive Regulation during Decision Making Shifts Behavioral Control between Ventromedial and Dorsolateral Prefrontal Value Systems, *The Journal of Neuroscience* 32(39), 13543–13554. http://doi.org/10.1523/JNEUROSCI.6387-11.2012

Inesi, M. E., Botti, S., Dubois, D., Rucker, D. D., & Galinsky, A. D. (2011). Power and Choice : Their Dynamic Interplay in Quenching the Thirst for Personal Control, *Psychological Science* (June). http://doi.org/10.1177/0956797611413936

Izuma, K., Saito, D. N., & Sadato, N. (2008). Article Processing of Social and Monetary Rewards in the Human Striatum, *Neuron.* 2008 Apr 24;58(2):284-94. doi: 10.1016/j.neuron.2008.03.020.

Jabbi, M., Bastiaansen, J., & Keysers, C. (2008). A Common Anterior Insula Representation of Disgust Observation , Experience and Imagination Shows Divergent Functional Connectivity Pathways, PLoS One. 2008 Aug 13;3(8):e2939. doi: 10.1371/journal.pone.0002939.

Judge, T. A., & Bono, J. E. (2001). Relationship of Core Self-Evaluations Traits — Self-Esteem , Generalized Self-Efficacy , Locus of Control , and Emotional Stability — With Job Satisfaction and Job Performance : A Meta-Analysis, *J Appl Psychol.* 2001 Feb;86(1):80-92. http://doi.org/10.1037//0021-9010.86.1.80

Juvonen, J., & Gross, E. F. (2005). The Rejected and the Bullied: Lessons About Social Misfits from Developmental Psychology. In *7th Annual Sydney Symposium of Social Psychology:" The Social Outcast: Ostracism, Social Exclusion, Rejection, and Bullying.", 7, Sydney, Australia;*. Psychology Press.

Kacou, A., & Law, G. P. I. (2012). Five arguments on the rationality of suicide terrorists. *Aggression and Violent Behavior*, Vol. 18, No. 5, pp. 539-547.

Kahane, G., Wiech, K., Shackel, N., Farias, M., Savulescu, J., & Tracey, I. (2012). The neural basis of intuitive and counterintuitive moral judgment. *Soc Cogn Affect Neurosci.* 2012 Apr;7(4):393-402. doi: 10.1093/scan/nsr005.

Kahnt, T., Park, S. Q., Haynes, J., & Tobler, P. N. (2014). Disentangling neural representations of value and salience in the human brain Disentangling neural representations of value and salience in the human brain, *Proc Natl Acad Sci U S A.* 2014 Apr 1;111(13):5000-5. doi: 10.1073/pnas.1320189111.

Karsh, N., & Eitam, B. (2015). I control therefore I do : Judgments of agency influence action selection. *Cognition*, *138*, 122–131. http://doi.org/10.1016/j.cognition.2015.02.002

Kay, A.C., Gaucher, D., McGregor, I. & Nash, K. (2010) Religious Belief as Compensatory Control *Pers Soc Psychol Rev February 2010 14: 37-48*

Kiehl, K. A., Smith, A. M., Hare, R. D., Mendrek, A., Forster, B. B., Brink, J., & Liddle, P. F. (2001). Limbic Abnormalities in Affective Processing by Criminal Psychopaths as Revealed by Functional Magnetic Resonance Imaging. *Biol Psychiatry*. 2001 Nov 1;50(9):677-84.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, 36(14), 1-16.

Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A., & Fernández, G. (2009). Reinforcement Learning Signal Predicts Social Conformity. *Neuron*, *61*(1), 140–151. http://doi.org/10.1016/j.neuron.2008.11.027

Knutson, B., & Wimmer, G. E. (2007). Splitting the Difference How Does the Brain Code Reward Episodes ?, *Ann N Y Acad Sci.* 2007 May;1104:54-69. http://doi.org/10.1196/annals.1390.020

Koechlin, E. (2003). The Architecture of Cognitive Control in the Human Prefrontal Cortex, *Science.* 2003 Nov 14;302(5648):1181-5.

Koenigs, M., & Tranel, D. (2008). Prefrontal Damage : Evidence from the Ultimatum Game, *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(4), 951–956. http://doi.org/10.1523/JNEUROSCI.4606-06.2007

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2008). Damage to the prefrontal cortex increases utilitarian moral judgements, *Nature.* 2007 Apr 19;446(7138):908-11.

Koolhaas, J. M., Bartolomucci, A., Buwalda, B., Boer, S. F. De, Flügge, G., Korte, S. M., … Fuchs, E. (2011). Neuroscience and Biobehavioral Reviews Stress revisited : A critical evaluation of the stress concept. *Neuroscience and Biobehavioral Reviews*, *35*(5), 1291–1301. http://doi.org/10.1016/j.neubiorev.2011.02.003

Kranz, G. S., Kasper, S., & Lanzenberger, R. (2010). Review Reward And The Serotonergic System, *Neuroscience.* 2010 Apr 14;166(4):1023-35. doi: 10.1016/j.neuroscience.2010.01.036.

Kriegeskorte, N., Formisano, E., Sorger, B., & Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proc Natl Acad Sci U S A.* 2007 Dec 18;104(51):20600-5.

Kriesberg, Louis. "Identity Issues." *Beyond Intractability*. Ed. Guy Burgess and Heidi Burgess. Conflict Research Consortium, University of Colorado, Boulder, Colorado, USA. July 2003. <http://www.beyondintractability.org/essay/identity-issues>.

Krueger, A. B. & Maleckova J. (2002). Education, Poverty, Political Violence And Terrorism:

Is There A Causal Connection? NBER Working Paper No. 9074

Lee, E. M., Klement, K. R., Ambler, J. K., Loewald, T., Comber, M., Hanson, S. A., … Sagarin, B. J. (2016). Altered States of Consciousness during an Extreme Ritual, *PLoS One*. 2016 May 13;11(5):e0153126. doi: 10.1371/journal.pone.0153126.

Leotti, L. A., & Ochsner, K. N. (2011). Born to Choose: The Origins and Value of the Need for Control, *Trends Cogn Sci*. 2010 October ; 14(10): 457–463. doi:10.1016/j.tics.2010.08.001.

Li, Q., Qin, S., Rao, L., Zhang, W., Ying, X., Guo, X., & Guo, C. (2011). Can Sophie ' s Choice Be Adequately Captured by Cold Computation of Minimizing Losses ? An fMRI Study of Vital Loss Decisions, *PLoS One*, *6*(3), 1–9. http://doi.org/10.1371/journal.pone.0017544

Li, W., Mai, X., Liu, C., & Moran, J. (2014). The default mode network and social understanding of others : what do brain connectivity studies tell us, *Frontiers in Human Neuroscience* (February), 1–15. http://doi.org/10.3389/fnhum.2014.00074

Li, Y., Vanni-mercier, Ã. G., & Isnard, J. (2016). The neural dynamics of reward value and risk coding in the human orbitofrontal cortex, *Brain*, 1–15. http://doi.org/10.1093/awwxxx

Lieberman, M. D. (2007). Social Cognitive Neuroscience : A Review of Core Processes. *Annu. Rev. Psychol. 2007. 58:259–89* http://doi.org/10.1146/annurev.psych.58.110405.085654

Longe, O., Maratos, F. A., Gilbert, P., Evans, G., Volker, F., Rockliff, H., & Rippon, G. (2010). NeuroImage Having a word with yourself : Neural correlates of self-criticism and self-reassurance. *NeuroImage*, *49*(2), 1849–1856. http://doi.org/10.1016/j.neuroimage.2009.09.019

MacDonald, D.A.; Holland, D. (2002). "Examination of the psychometric properties of the temperament and character inventory self-transcendence dimension". *Personality and Individual Differences* 32 (6): 1013–1027. doi:10.1016/S0191-8869(01)00107-6.

Motzkin, J.C., Newman, J.P., Kiehl, K.A., & Koenigs, M. (2012). Reduced Prefrontal Connectivity in Psychopathy, *J Neurosci*. 2011 November 30; 31(48): 17348–17357. doi:10.1523/JNEUROSCI.4215-11.2011..

Herringa, R., Phillips, M., Almeida, J., Insana, S., & Germain, A. (2013). Post-traumatic stress symptoms correlate with smaller subgenual cingulate, caudate, and insula volumes in unmedicated combat veterans. *Psychiatry Res*. 2012 August ; 203(2-3): 139–145. doi:10.1016/j.pscychresns.2012.02.005.

Mason, M. F., Dyer, R., & Norton, M. I. (2009). Organizational Behavior and Human Decision Processes Neural mechanisms of social influence. *Organizational Behavior and Human Decision Processes*, *110*(2), 152–159. http://doi.org/10.1016/j.obhdp.2009.04.001

Mcclure, S. M., & Cohen, J. D. (2012). Separate Neural Systems Value Immediate and Delayed Monetary Rewards, *Science* 503(2004). http://doi.org/10.1126/science.1100907

McGraw, a. P., & Tetlock, P. E. (2005). Taboo Trade-Offs, Relational Framing, and the Acceptability of Exchanges. *Journal of Consumer Psychology*, *15*(1), 2–15. http://doi.org/10.1207/s15327663jcp1501_2

Merari, A.; Diamant, I.; Bibi, A.; Broshi, Y.; Zakin, G. (2010) Personality characteristics of "self martyrs"/"suicide bombers" and organizers of suicide attacks. *Terrorism and Political Violence*, Vol 22(1), Jan 2010, 87-101. http://dx.doi.org/10.1080/09546550903409312

Mier, D., Kirsch, P., & Meyer-lindenberg, A. (2010). Neural substrates of pleiotropic action of genetic variation in COMT: A meta-analysis, *Mol Psychiatry.* 2010 Sep;15(9):918-27. http://doi.org/10.1038/mp.2009.36

Mitchell, J. P., Macrae, C. N., Banaji, M. R., & Hall, W. J. (2006). Dissociable Medial Prefrontal Contributions to Judgments of Similar and Dissimilar Others, Neuron, May 18;50(4):655-63. http://doi.org/10.1016/j.neuron.2006.03.040

Moll, J., Oliveira-souza, R. De, Bramati, I. E., & Grafman, J. (2002). Functional Networks in Emotional Moral and Nonmoral Social Judgments, *Neuroimage.* 2002 Jul;16(3 Pt 1):696-703. http://doi.org/10.1006/nimg.2002.1118

Morgan, T. J. H., & Laland, K. N. (2012). The biological bases of conformity, *Frontiers in Neuroscience*, *6*, 87. http://doi.org/10.3389/fnins.2012.00087

Moskowitz, G. B., Favoritism, I., Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective taking : Decreasing stereotype accessibility and in-group favoritism Perspective-Taking : Decreasing Stereotype Expression , *J Pers Soc Psychol.* 2000 Apr;78(4):708-24. http://doi.org/10.1037//0022-3514.78.4.708

Norenzayan, A., Shariff, A. F., Willard, A. K., Slingerland, E., Will, M., Mcnamara, R. A., & Henrich, J. (2014). The Cultural Evolution of Prosocial Religions, *Behav Brain Sci.* 2016 Jan;39:e1. doi: 10.1017/S0140525X14001356.

Norris, P., & Inglehart, R. (2004). Sacred and Secular : Religion and Politics Worldwide, *Cambridge University Press.*

Northoff, G., Heinzel, A., Greck, M. De, Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain — A meta-analysis of imaging studies on the self, *Neuroimage.* 2006 May 15;31(1):440-57. http://doi.org/10.1016/j.neuroimage.2005.12.002

Obhi, S. S., Swiderski, K. M., & Brubacher, S. P. (2012). Induced power changes the sense of agency. *Consciousness and Cognition*, *21*(3), 1547–1550. http://doi.org/10.1016/j.concog.2012.06.008

Ochsner, K. N., Bunge, S. A., Gross, J. J., & Gabrieli, J. D. E. (2002). Rethinking Feelings : An fMRI Study of the Cognitive Regulation of Emotion, *J Cogn Neurosci.* 2002 Nov 15;14(8):1215-29.

Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion, *Trends Cogn Sci.* 2005 May;9(5):242-9. http://doi.org/10.1016/j.tics.2005.03.010

Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D. E., & Gross, J. J. (2004). For better or for worse : neural systems supporting the cognitive down- and up-regulation of negative emotion, *Neuroimage.* 2004 Oct;23(2):483-99. http://doi.org/10.1016/j.neuroimage.2004.06.030

Olsson, A., & Ochsner, K. N. (2007). The role of social cognition in emotion, *Trends Cogn Sci.* 2008 Feb;12(2):65-71. doi: 10.1016/j.tics.2007.11.010.

Öngür, D., & Price, J. L. (2000). The Organization of Networks within the Orbital and Medial Prefrontal Cortex of Rats , Monkeys and Humans, *Cereb Cortex.* 2000 Mar;10(3):206-19.

Paladino, M., Rodriguez, R., Rodriguez, A., Gaunt, R., & Demoulin, S. (2002). Differential Association of Uniquely and Non Uniquely Human Emotions with the Ingroup and the Outgroup, *Group Processes & Intergroup Relations* 2002 Vol 5(2) 105–117.

Passingham, R. E., Bengtsson, S. L., & Lau, H. C. (2009). Medial frontal cortex : from self-generated action to reflection on one ' s own performance, *Trends Cogn Sci.* 2010 Jan;14(1):16-21. doi: 10.1016/j.tics.2009.11.001

Pettigrew, T. F., & Tropp, L. R. (2006). Interpersonal Relations And Group Processes: A Meta-Analytic Test of Intergroup Contact Theory, *J Pers Soc Psychol.* 2006 May;90(5):751-83.

Pincus, M., LaViers, L., Prietula, M. J., & Berns, G. (2014). The conforming brain and deontological resolve. *PloS One*, *9*(8), e106061. http://doi.org/10.1371/journal.pone.0106061

Preston, J. L., & Ritter, R. S. (2013). Different Effects of Religion and God on Prosociality With the Ingroup and Outgroup. *Pers Soc Psychol Bull* August 22, 2013 0146167213499937

Rai, M. (2004). Hindu Rulers, Muslim Subjects: Islam, Rights, and the History of Kashmir. Princeton: Princeton University Press.

Raine A. (2000). Reduced Prefrontal Gray Matter Volume and Reduced Autonomic Activity in Antisocial Personality Disorder, *Arch Gen Psychiatry.* 2000 Feb;57(2):119-27.

Ramnani, N., & Owen, A. M. (2004). Anterior Prefrontal Cortex : Insights Into Function From Anatomy And Neuroimaging, *Nat Rev Neurosci.* 2004 Mar;5(3):184-94. http://doi.org/10.1038/nrn1343

Reddish, P., Fischer, R., & Bulbulia, J. (2013). Let ' s Dance Together : Synchrony , Shared Intentionality and Cooperation, *PLoS One.* 2013 Aug 7;8(8):e71182. doi: 10.1371/journal.pone.0071182.

Rice C., quoted in Munayer Salim J, Loden Lisa, Through My Enemy's Eyes: Envisioning Reconciliation in Israel-Palestine (2014), quote: "The Palestinian-Israeli divide may be the most intractable conflict of our time."

Ridderinkhof, K. R. (2008). The Role of the Medial Frontal Cortex in Cognitive Control, *Science.* 2004 Oct 15;306(5695):443-7.http://doi.org/10.1126/science.1100301

Rilling, J. K., King-casas, B., & Sanfey, A. G. (2008). The neurobiology of social decision-making. *Curr Opin Neurobiol.* 2008 Apr;18(2):159-65. doi: 10.1016/j.conb.2008.06.003.

Rochat, P., Dias, M. D. G., Broesch, T., Passos-ferreira, C., & Berg, B. (2009). Fairness in Distributive Justice by 3- and 5-Year-Olds Across Seven Cultures, *Journal of Cross-Cultural Psychology* May 2009 vol. 40 no. 3 416-442

Rolls, E. T. (2000). The Orbitofrontal Cortex and Reward, *Cereb Cortex.* 2000 Mar;10(3):284-94.

Rolls, E. T. (2015). Progress in Neurobiology Taste , olfactory , and food reward value processing in the brain. *Progress in Neurobiology*, *127-128*, 64–90. http://doi.org/10.1016/j.pneurobio.2015.03.002

Rotter, J. (1966) 'Generalized expectancies for internal versus external control of reinforcement', *Psychological Monographs* 80 (1), 1–28);

Ruby, P., Decety, J., Inserm, U., Thomas, A., & Cedex, L. (2003). What you believe versus what you think they believe : a neuroimaging study of conceptual perspective-taking, *Eur J Neurosci.* 2003 Jun;17(11):2475-80. http://doi.org/10.1046/j.1460-9568.2003.02673.x

Rudebeck, P. H., Behrens, T. E., Kennerley, S. W., Baxter, M. G., Buckley, M. J., Walton, M. E., & Rushworth, M. F. S. (2008). Frontal Cortex Subregions Play Distinct Roles in Choices between Actions and Stimuli, *J Neurosci.* 2008 Dec 17;28(51):13775-85. http://doi.org/10.1523/JNEUROSCI.3541-08.2008

Ryan, M. E., & Francis, A. J. P. (2012). Locus of Control Beliefs Mediate the Relationship Between Religious Functioning and Psychological Health, *J Relig Health.* 2012 Sep;51(3):774-85. doi: 10.1007/s10943-010-9386-z.

Sachdeva, S., & Medin, D. (2009). Group identity salience in sacred value based cultural conflict: an examination of the Hindu-Muslim identities in the Kashmir and Babri mosque issues. *In Proceedings of the 31st Annual Conference of the Cognitive Science Society*, 3111–3114. Retrieved from http://csjarchive.cogsci.rpi.edu/Proceedings/2009/papers/677/paper677.pdf

Sakagami, M., & Pan, X. (2007). Functional role of the ventrolateral prefrontal cortex in decision making, *Curr Opin Neurobiol.* 2007 Apr;17(2):228-33 http://doi.org/10.1016/j.conb.2007.02.008

Sanfey, A. (2007) Social Decision-Making: Insights from Game Theory and Neuroscience, Science Vol. 318, 598-602.

Sanfey, A.G., Rilling J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D. (2003). The Neural Basis of Economic Decision-Making in the, *Science* Vol.300 (June), 1755–1759.

Sawamoto, N., Honda, M., Okada, T., Hanakawa, T., Kanda, M., Fukuyama, H., … Shibasaki, H. (2000). Expectation of Pain Enhances Responses to Nonpainful Somatosensory Stimulation in the Anterior Cingulate Cortex and Parietal Operculum / Posterior Insula : an Event-Related Functional Magnetic Resonance Imaging Study, *J Neurosci.* 2000 Oct 1;20(19):7438-45.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people The role of the temporo-parietal junction in " theory of mind " *Neuroimage.* 2003 Aug;19(4):1835-42. http://doi.org/10.1016/S1053-8119(03)00230-1

Schmitz, T. W., & Johnson, S. C. (2006). Self-appraisal decisions evoke dissociated dorsal — ventral aMPFC networks, *Neuroimage.* 2006 Apr 15;30(3):1050-8. http://doi.org/10.1016/j.neuroimage.2005.10.030

Schmitz, T. W., & Johnson, S. C. (2007). Relevance to self : A brief review and framework of neural systems underlying appraisal, *Neurosci Biobehav Rev.* 2007;31(4):585-96. http://doi.org/10.1016/j.neubiorev.2006.12.003

Schultz, Norman. "Distinguishing facts from values" *Beyond Intractability.* Ed. Guy Burgess and Heidi Burgess. Conflict Research Consortium, University of Colorado, Boulder, Colorado, USA. Nov 2003. <http://www.beyondintractability.org/essay/facts-values>.

Seymour, B., Doherty, J. P. O., Dayan, P., Koltzenburg, M., Jones, A. K., Dolan, R. J., … Frackowiak, R. S. (2004). Temporal difference models describe higher-order learning in humans, *Nature* 429(June), 664–667. http://doi.org/10.1038/nature02636.1.

Sharot, T., Kanai, R., Marston, D., Korn, C. W., Rees, G., & Dolan, R. J. (2012). Selectively altering belief formation in the human brain, *Proc Natl Acad Sci U S A* 109(42). http://doi.org/10.1073/pnas.1205828109

Sheikh, H., Ginges, J., & Atran, S. (2013). Sacred values in the Israeli–Palestinian conflict: resistance to social influence, temporal discounting, and exit strategies. *Annals of the New York Academy of Sciences*, *1299*, 11–24. http://doi.org/10.1111/nyas.12275

Sheikh, H., Gómez, Á., & Atran, S. (2016). Empirical Evidence for the Devoted Actor Model, *Current Anthropology* 57, no. S13 (June 2016): S204-S209. http://doi.org/10.1086/686221

Shih, J.C., Chen, K., Ridd, M.J. (2010) MONOAMINE OXIDASE: From Genes to Behavior *Annu Rev Neurosci*. 1999 ; 22: 197–217. doi:10.1146/annurev.neuro.22.1.197.

Sosis, R., & Bressler, E. R. (2016). Cooperation and Commune Longevity : A Test of the Costly Signaling Theory of Religion, *Cross-Cultural Research May 2003 vol. 37 no. 2 211-239* http://doi.org/10.1177/1069397103251426

Sosis, R., & Ruffle, B.J. (2003) Religious Ritual and Cooperation: Testing for a Relationship on Israeli Religious and Secular Kibbutzim, *Current Anthropology*, Vol. 44, No. 5 (December 2003), pp. 713-722

Sosis, R., Kress, H. C., & Boster, J. S. (2007). Scars for war : evaluating alternative signaling explanations for cross-cultural variance in ritual costs, *Evolution and Human Behavior* Vol. 28, 234–247. http://doi.org/10.1016/j.evolhumbehav.2007.02.007

Summer, J. J., Hassabis, D., & Maguire, E. A. (2009). Cortical midline involvement in autobiographical memory, *NeuroImage* Vol. 44, 1188–1200. http://doi.org/10.1016/j.neuroimage.2008.09.033

Svingos, A. L., & Garzo, M. (2001). Mu -Opioid Receptors in the Ventral Tegmental Area Are Targeted to Presynaptically and Directly Modulate Mesocortical Projection Neurons, *Synapse.* 2001 Sep 1;41(3):221-9.

Swann WB, Jr, Pelham BW, Chidester TR (1988) Change through paradox: Using self-verification to alter beliefs. *J Pers Soc Psychol* 54(2):268–273.

Swann, W. B., Gómez, A., Seyle, D. C., Morales, J. F., & Huici, C. (2009). Identity fusion: the interplay of personal and social identities in extreme group behavior. *Journal of Personality and Social Psychology*, *96*(5), 995–1011. http://doi.org/10.1037/a0013668

Swann, W. B., Jetten, J., Gómez, Á., Whitehouse, H., & Bastian, B. (2012). When group membership gets personal: A theory of identity fusion. *Psychological Review*, *119*(3), 441–456. http://doi.org/10.1037/a0028589

Tajfel, H., & Turner, J. C. (1979). "An integrative theory of intergroup conflict". In W. G. Austin & S. Worchel. *The social psychology of intergroup relations*. Monterey, CA: Brooks/Cole. pp. 33–47.

Tan, E., Lim, E. C. P., Teo, Y., Lim, Y., & Sia, A. T. (2009). Ethnicity and OPRM variant independently predict pain perception and patient-controlled analgesia usage for post-operative pain, *Molecular Pain* 2009 5:32. http://doi.org/10.1186/1744-8069-5-32

Tanaka, S. C., Balleine, B. W., & Doherty, J. P. O. (2008). Calculating Consequences : Brain Systems That Encode the Causal Effects of Actions, *The Journal of Neuroscience : The*

Official Journal of the Society for Neuroscience, *28*(26), 6750–6755. http://doi.org/10.1523/JNEUROSCI.1808-08.2008

Tetlock, P. E. (2003). Thinking the unthinkable: sacred values and taboo cognitions. *Trends in Cognitive Sciences*, *7*(7), 320–324. http://doi.org/10.1016/S1364-6613(03)00135-9

Tetlock, P. E., Kristel, O. V, Elson, S. B., Green, M. C., & Lerner, J. S. (2000). The Psychology of the Unthinkable : Taboo Trade-Offs , Forbidden Base Rates , and Heretical Counterfactuals, J *Pers Soc Psychol.* 2000 May;78(5):853-70. http://doi.org/10.1037//0022-3514.78.5.853

Thomson, J. J. (1985). The Trolley Problem, *The Yale Law Journal*, Vol. *94*, No. *6* (May, 1985), pp. 1395-1415.

Turner, J.; Oakes, P. (1986). "The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence". *British Journal of Social Psychology* **25** (3): 237–252. doi:10.1111/j.2044-8309.1986.tb00732.

Utevsky, A. V, Smith, D. V, & Huettel, S. A. (2014). Precuneus Is a Functional Core of the Default-Mode Network, *J Neurosci.* 2014 Jan 15;34(3):932-40. doi: 10.1523/JNEUROSCI.4227-13.2014.

Vaes, J., Leyens, J., Paladino, M. P., & Miranda, M. P. (2012). We are human , they are not : Driving forces behind outgroup dehumanisation and the humanisation of the ingroup, *European Review of Social Psychology*, 23:1, 64-106.

Visser, M., Embleton, K. V, Jefferies, E., Parker, G. J., & Ralph, M. A. L. (2010). Neuropsychologia The inferior , anterior temporal lobes and semantic memory clarified : Novel evidence from distortion-corrected fMRI. *Neuropsychologia*, *48*(6), 1689–1696. http://doi.org/10.1016/j.neuropsychologia.2010.02.016

Vogeley, K., May, M., Ritzl, A., Falkai, P., Zilles, K., & Fink, G. R. (2004). Neural Correlates of First-Person Perspective as One Constituent of Human Self-Consciousness, *J Cogn Neurosci.* 2004 Jun;16(5):817-27.

Volkow, N. D., Wang, G., Fowler, J. S., Logan, J., Gerasimov, M., Maynard, L., … Franceschi, D. (2001). Therapeutic Doses of Oral Methylphenidate Significantly Increase Extracellular Dopamine in the Human Brain, *J Neurosci.* 2001 Jan 15;21(2):RC121.

Vollenweider, F. X., Vontobel, P., Ph, D., Hell, D., & Leenders, K. L. (1998). 5-HT Modulation of Dopamine Release in Basal Ganglia in Psilocybin-Induced Psychosis in Man — A PET Study with [11 C] raclopride, *Neuropsychopharmacology.* 1999 May;20(5):424-33.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory : A meta-analysis, *Cogn Affect Behav Neurosci*. 2003 Dec;3(4):255-74.

Way, B. M., Taylor, S. E., & Eisenberger, N. I. (2009). Variation in the Mu -opioid receptor gene (OPRM1) is associated with dispositional and neural sensitivity to social rejection, *Proc Natl Acad Sci U S A 106*(35), 15079–15084.

Weems, C. F., Silverman, W. K., Rapee, R. M., & Pina, A. A. (2003). The Role of Control in Childhood Anxiety Disorders, *Cognitive Therapy and Research 27*(5), 557–568.

Wheeler, M. E., & Buckner, R. L. (2004). Functional-anatomic correlates of remembering and knowing, *Neuroimage.* 2004 Apr;21(4):1337-49. http://doi.org/10.1016/j.neuroimage.2003.11.001

Whitson, J. A., & Galinsky, A. D. (2008). Lacking Control Increases Illusory Pattern Perception, *Science.* 2008 Oct 3;322(5898):115-7. doi: 10.1126/science.1159845.

Wicker, B., Keysers, C., Plailly, J., Royet, J., Gallese, V., Rizzolatti, G., … Garnier, A. T. (2003). Both of Us Disgusted in My Insula : The Common Neural Basis of Seeing and Feeling Disgust, *Neuron.* 2003 Oct 30;40(3):655-64.

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and Cooperation, *Association for Psychological Science*, 1–5.

Wisniewski, X. D., Reverberi, C., Momennejad, I., Kahnt, X., & Haynes, J. (2015). The Role of the Parietal Cortex in the Representation of Task – Reward Associations, *J Neurosci.* 2015 Sep 9;35(36):12355-65. doi: 10.1523/JNEUROSCI.4882-14.2015.

Xygalatas, D., Mitkidis, P., Fischer, R., Reddish, P., Skewes, J., Geertz, A. W., … Bulbulia, J. (2016). Extreme Rituals Promote Prosociality, *Psychol Sci.* 2013 Aug;24(8):1602-5. doi: 10.1177/0956797612472910. http://doi.org/10.1177/0956797612472910

Yamada, M., Uddin, L. Q., Takahashi, H., Kimura, Y., Takahata, K., & Kousa, R. (2013). Superiority illusion arises from resting-state brain networks modulated by dopamine, *Proc Natl Acad Sci U S A* 11(11). http://doi.org/10.1073/pnas.1221681110

Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment, *Proc Natl Acad Sci U S A.* 2007 May 15;104(20):8235-40.

Young, L., & Koenigs, M. (2007). Investigating emotion in moral cognition : a review of evidence from functional neuroimaging and neuropsychology, *Br Med Bull.* 2007;84:69-79. http://doi.org/10.1093/bmb/ldm031

Zahn, R., Moll, J., Paiva, M., Garrido, G., Krueger, F., Huey, E. D., & Grafman, J. (2009). The Neural Basis of Human Social Values : Evidence from Functional MRI, *Cereb Cortex.* 2009 Feb;19(2):276-83. doi: 10.1093/cercor/bhn080

Zaki, J., Schirmer, J., & Mitchell, J. P. (2011). Social Influence Modulates the Neural Computation of Value. *Psychological Science*, *22*(7), 894–900. http://doi.org/10.1177/0956797611411057

Annex

Field survey for the selection of participants and their sacred values:

Please remember: All answers are completely confidential. Please answer honestly.

Demographics

Thank you for participating in our study. Please respond to the following demographic questions for statistical reasons.1. Age:

2. Marital Status
a. Single
b. Engaged
c. Married
d. Divorced
e. Widowed

3. Highest Education
a. Primary
b. Middle level
c. Secondary/Matric/Diploma
d. FA, FSc, I.Com/Certificate or equivalent
e. Bachelor's Degree
f. Master's Degree
g. Other _____

4. How does your income level compare the national average income of Spain (approx. €24,200)?
a. Well below average
b. Below average
c. About average
d. Above average
e. Well above average

5. How important are religious beliefs in your life?
1 Not at all
2
3
4
5
6
7 The most important thing in my life

6. How often do you pray in private?
a. Five times a day
b. Once or twice a day
c. Once or twice a week
d. Once or twice a month
e. Once or twice a year
f. Never

7. How often do you attend the mosque?
a. Five times a day
b. Once or twice a day
c. Once or twice a week
d. Once or twice a month
e. Once or twice a year
f. Never

8. Which Islamic jamaat do you most affiliate with?

9. Which sect of Islam do you belong to?
a. Sunni
b. Shia
c. Deobandi
d. Ahle Hadee

10. Please rank which of the following two items in terms of which guide your behavior more:

a. _____ Avoiding Hell fire
b. _____Entering Heaven

11. Please answer the following question regarding your preferred type of government in an ideal world:

	Very Secular	Somewhat Secular	Neither Secular nor Religious	Somewhat Religious	Very Religious
What is your preferred type of government?	1	2	3	4	5

12. Please answer the following question about your political preference about European politics. Very liberal corresponds to supporting social democratic political parties, and very conservative corresponds to supporting Christian conservative parties:

	Very Liberal	Somewhat Liberal	Neither Liberal nor Conservative	Somewhat Conservative	Very Conservative
What is your political affiliation in European politcs?	1	2	3	4	5

13. Here is a list of statements some Muslims agree with and others disagree with. Please tell us how much you personally agree or disagree with each statement.	*Strongly disagree*	*Somewhat disagree*	*Neither disagree nor agree*	*Somewhat agree*	*Strongly agree*
1. Our existence as Muslims is under constant threat.	1	2	3	4	5
2. Western nations seek to manipulate and take advantage of Muslim nations.	1	2	3	4	5
3. There is a war against Islam by Western nations.	1	2	3	4	5
4. If Muslims had a major say in the world, the world would be a much better place.	1	2	3	4	5
5. Muslims deserve special treatment.	1	2	3	4	5
6. It really makes me angry when others criticize Muslims.	1	2	3	4	5
7. I will never be satisfied until Muslims get the recognition they deserve.	1	2	3	4	5
8. The fighting of the Taliban, Al Qaida, ISIS is justified.	1	2	3	4	5
9. The violence of Hamas is justified	1	2	3	4	5
10. Groups like Taliban, Al Qaida, ISIS are created and currently funded by Mossad, RAW, or CIA to divide and conquer Muslim countries.	1	2	3	4	5
11. The violence of Pakistani mujahedeen is justified.	1	2	3	4	5
12. Muslims everywhere should try to restore the Caliphate (borderless/federation) by all means possible.	1	2	3	4	5
13. Muslims must unite in order to oppose Western forces by all means necessary.	1	2	3	4	5
14. Spreading Islam with force in every part of the world can be an act of justifiable Jihad.	1	2	3	4	5
15. The strictest forms of sharia should apply to both Muslims and non-Muslims.	1	2	3	4	5
16. Theft should be punishable by cutting off of hands.	1	2	3	4	5
17. Adultery should be punishable by stoning or whipping.	1	2	3	4	5
18. Apostasy should be punishable by the death penalty.	1	2	3	4	5
19. Sharia should play a much larger role in Pakistani law.	1	2	3	4	5
20. Sharia should play a much larger role in Spanish law.	1	2	3	4	5
21. The killing of governor of Punjab, Salman Taseer, was justified because of his defense of a Christian woman, Asia Bibi, who was sentenced to death for allegedly insulting the prophet.	1	2	3	4	5
22. Sunni/Shia/Deobandi/Adle Hadees are not true Muslims.	1	2	3	4	5

23. The killing of someone who mocks the prophet is justified	1	2	3	4	5
24. Violence against other Muslims is never justified.	1	2	3	4	5

14. Here is a list of groups or parties in Pakistan and the Middle East. Some people hold very favorable views of each group and others hold unfavorable opinions of each group. Please tell us if you have a favorable or unfavorable opinion about each of the following groups or parties?

	Very Unfavorable	*Somewhat Unfavorable*	*Neither/ Don't know*	*Somewhat Favorable*	*Very Favorable*
Pakistan People's Party (PPP)	1	2	3	4	5
Muttahida Majlis-e-Amal (MMA)	1	2	3	4	5
Pakistan Muslim League (PML-N)	1	2	3	4	5
Pakistan Muslim League (PML-Q)	1	2	3	4	5
Tehrik-i-Taliban Pakistan (TTP)	1	2	3	4	5
Lashkar-i-Tayyiba/Jamaat-ud-Dawa	1	2	3	4	5
Ad-dawlah al-islamiyah	1	2	3	4	5
Jamaat e-Islami	1	2	3	4	5
Tahrik-i-insaf	1	2	3	4	5
Muslim Brotherhood	1	2	3	4	5

15. People who join jihadi groups (Taliban, Al Qaida, ISIS) do it primarily for:
A) Money
B) Because they sincerely believe in the cause of these groups
C) Both
D) Other: _____

Kashmir

16) I **agree/disagree** that India should have no right to ownership of Kashmir.

Please respond to one of the following scenarios based on whether you agree or disagree with the above issue.

Agree:

16.a.1) If instead of allowing Kashmir to belong to Pakistan or become independent the Indian government gave large amounts of tax refunds, local aid, heavy economic investment (i.e. more jobs), and other monetary compensation to all Kashmiri families so they could live a higher material quality of life, or any other **purely financial** incentives, to stay part of India, would you find it acceptable?

Yes Maybe No

16.b.1) What if Kashmiris wanted to stay part of India, would that be acceptable?

Yes Maybe No

Disagree:
16.a.2) If allowing the whole of Kashmir to belong to Pakistan came with an agreement from the Pakistani government that it would give large amounts of tax refunds, local aid, heavy economic investment (i.e. more jobs), and other monetary compensation to all Kashmiri families so they could live a higher material quality of life, or any **purely financial** reasons, than if they became independent or stayed part of India, would you find it acceptable?

Yes Maybe No

16b.2) What if Kashmiris wanted to belong entirely to Pakistan, would that be acceptable?

Yes Maybe No

16.c) Please select all the actions that you would be willing to take to defend your position on this issue.
1 - Nothing at all
2 - Persuade people one-on-one
3 - Protest or apply public pressure for governments to find a diplomatic solution
4 - Protest or apply public pressure for governments to find a military solution (if issue in other country)/violently protest (if issue in same country)
5 - Support in spirit or financially a non-state militant group
6 - Join a non-state militant group
7 - Carry out militant actions to fight and die on my own, even if I had no group or support network

<u>Gay Marriage</u>
17) I **agree/disagree** that marriage between homosexual partners should be illegal in Muslim countries.

Please respond to one of the following scenarios based on whether you agree or disagree with the above issue.

Agree:
17.a.1) If a Muslim country legalized homosexual marriage because doing so led to a beneficial relationship of economic benefits from non-Muslim countries such as increased foreign aid, more jobs due to corporate partnerships, profit from the trade-relations etc., or any other **purely financial** reasons, would you find it acceptable?

Yes Maybe No

17.b.1) If the people and government of a Muslim country decided they want to legalize homosexual marriage, would that be acceptable?

Yes Maybe No

Disagree:
17.a.2) If a Muslim country did not legalize homosexual marriage because doing so led to a beneficial relationship of economic benefits from Muslim countries such as increased foreign aid, more jobs due to corporate partnerships, profit from the trade-relations etc., or any other **purely financial** reasons, would you find it acceptable?

Yes Maybe No

17.b.2) If the people and government of a Muslim country decided they did not want to legalize homosexual marriage, would that be acceptable?

Yes Maybe No

17.c) Please select all the actions that you would be willing to take to defend your position on this issue.

1 - Nothing at all
2 - Persuade people one-on-one
3 - Protest or apply public pressure for governments to find a diplomatic solution
4 - Protest or apply public pressure for governments to find a military solution (if issue in other country)/violently protest (if issue in same country)
5 - Support in spirit or financially a non-state militant group
6 - Join a non-state militant group
7 - Carry out militant actions to fight and die on my own, even if I had no group or support network.

U.S./Western Military
18) I **agree/disagree** that all US and other Western military forces should be expelled from all Muslim lands.

Please respond to one of the following scenarios based on whether you agree or disagree with the above issue.

Agree:
18.a.1) If Muslim countries allowed US and other Western military forces to stay on their lands because doing so led to a beneficial relationship of economic benefits such as increased foreign aid, more jobs due to corporate partnerships, profit from the trade-relations etc., or any other **purely financial** reasons, would you find it acceptable?

Yes Maybe No

18.b.1) If a Muslim country's people and government wanted US and other Western military forces on their lands, would that be acceptable?

Yes Maybe No

Disagree:
18.a.2) If Muslim countries did expel US and other Western military forces from their lands because doing so led to a beneficial relationship of economic benefits such as increased foreign aid, more jobs due to corporate partnerships, profit from the trade-relations etc., with other Muslim countries and their companies, or any other **purely financial** reasons, would you find it acceptable?

Yes Maybe No

18.b.2) If all Muslim countries' people and governments wanted to expel US and other Western military forces from their lands, would that be acceptable?

Yes Maybe No

18.c) Please select all the actions that you would be willing to take to defend your position on this issue.

1 - Nothing at all
2 - Persuade people one-on-one
3 - Protest or apply public pressure for governments to find a diplomatic solution
4 - Protest or apply public pressure for governments to find a military solution (if issue in other country)/violently protest (if issue in same country)
5 - Support in spirit or financially a non-state militant group
6 - Join a non-state militant group
7 - Carry out militant actions to fight and die on my own, even if I had no group or support network.

Drone Strikes
19) I **agree/disagree** that there should be an immediate stop to all U.S. drone strikes in the northwestern frontier provinces.

Please respond to one of the following scenarios based on whether you agree or disagree with the above issue.

Agree:
19.a.1) If Pakistan allowed U.S. drone strikes to continue because doing so led to a beneficial relationship of economic benefits such as increased foreign aid, more jobs due to corporate partnerships, profit from the trade-relations etc., or any other **purely financial** reasons, would you find it acceptable?

Yes Maybe No

19.b.1) If the Pakistani people and government wanted the US to continue its drone strikes in northwestern frontier because it was helping Pakistan achieve its military objectives, would that be acceptable?

Yes Maybe No

Disagree:
19.a.2) If Pakistan didn't allow U.S. drone strikes to continue because doing so led to a beneficial relationship of economic benefits such as increased foreign aid, more jobs due to corporate partnerships, profit from the trade-relations etc., with other Muslim countries and their companies, or any other **purely financial** reasons, would you find it acceptable?

Yes Maybe No

19.b.2) If the Pakistani people and government did not want the US to continue its drone strikes in northwestern frontier, would that be acceptable?

Yes Maybe No

19.c) Please select all the actions that you would be willing to take to defend your position on this issue.
1 - Nothing at all
2 - Persuade people one-on-one
3 - Protest or apply public pressure for governments to find a diplomatic solution
4 - Protest or apply public pressure for governments to find a military solution (if issue in other country)/violently protest (if issue in same country)
5 - Support in spirit or financially a non-state militant group
6 - Join a non-state militant group
7 - Carry out militant actions to fight and die on my own, even if I had no group or support network.

<u>Palestinian Right of Return</u>
20) I **agree/disagree** that Palestinians have a right to return to their homes in Israel.

Please respond to one of the following scenarios based on whether you agree or disagree with the above issue.

Agree:
20.a.1) If instead of allowing the Palestinians to return to their homes, the Israeli government offered large amounts of money to each family worth much more than their old homes to purchase new homes (or anything else) of much higher quality in Palestine, or any other **purely financial** incentives, would you find it acceptable?

Yes Maybe No

20.b.1) If the Palestinian people and government no longer wanted a right of return, would that be acceptable?

Yes Maybe No

Disagree:
20.a.2) If the Palestinians were given a right to return to their homes in Israel because doing so would improve trade-relations between Israel and other Arab countries leading to large financial gains for Israeli companies and the economy in general, or any other **purely financial** incentives, would you find it acceptable?

Yes Maybe No

20.b.2) If the Israeli people and government wanted to offer Palestinians a right of return, would that be acceptable?

Yes Maybe No

20.c) Please select all the actions that you would be willing to take to defend your position on this issue.

1 - Nothing at all
2 - Persuade people one-on-one
3 - Protest or apply public pressure for governments to find a diplomatic solution
4 - Protest or apply public pressure for governments to find a military solution (if issue in other country)/violently protest (if issue in same country)
5 - Support in spirit or financially a non-state militant group
6 - Join a non-state militant group
7 - Carry out militant actions to fight and die on my own, even if I had no group or support network.

<u>Palestinian Statehood</u>
21) I **agree/disagree** that there should be no Israel and only one Palestinian state from river to sea.

Please respond to one of the following scenarios based on whether you agree or disagree with the above issue.

Agree:
21.a.1) If instead a two-state solution was offered of a sovereign Palestine next to Israel which included large amounts of foreign aid from various governments, trade-agreements, corporate partnerships, and other monetary compensation to Palestine so Palestinians would have a strong economy and high material quality of life, would you find it acceptable?

Yes Maybe No

21.b.1) If the Palestinian people and government wanted a two-state solution with a sovereign Palestine and Israel side by side, would that be acceptable?

Yes Maybe No

Disagree:
21.a.2) If only one Palestine state from river to sea led to an increased economy for everyone as a result consolidated land usage, foreign aid from Muslim governments, trade-agreements, corporate partnerships, and other monetary compensation, so Muslims and Jews would have a higher material quality of life, would you find it acceptable?

Yes Maybe No

21.b.2) If the Israeli and Palestinian people and governments wanted a one-state solution with both people living in a single Palestine, would that be acceptable?

Yes Maybe No

21.c) Please select all the actions that you would be willing to take to defend your position on this issue.

1 - Nothing at all
2 - Persuade people one-on-one
3 - Protest or apply public pressure for governments to find a diplomatic solution
4 - Protest or apply public pressure for governments to find a military solution (if issue in other country)/violently protest (if issue in same country)
5 - Support in spirit or financially a non-state militant group
6 - Join a non-state militant group
7 - Carry out militant actions to fight and die on my own, even if I had no group or support network.

Halal Food in Public Centers

22) I **agree/disagree** that the Spanish government should make halal food available in public centers (schools, hospitals, shelters, prisons).

Please respond to one of the following scenarios based on whether you agree or disagree with the above issue.

Agree:
22.a.1) If instead of offering halal food in public centers the Spanish government offered large tax refunds to Muslims so that they could purchase halal food (or anything else) on their own with the money being returned, or any other **purely financial** incentives, would you find it acceptable?

Yes Maybe No

22.b.1) If the majority of Muslim people in Spain decided that they don't need the Spanish government to provide them with halal food in public centers, would that be okay?

Yes Maybe No

Disagree:
22.a.2) If the Spanish government decided that they would offer halal food in public centers because doing so would mean a revenue boost to local halal food companies which would translate into more jobs and revenue for the government through taxes, would you find it acceptable?

Yes Maybe No

22.b.2) If the majority of Spanish people and the government decided that they wanted to provide halal food in public centers, would that be okay?

Yes Maybe No

22.c) Please select all the actions that you would be willing to take to defend your position on this issue.

1 - Nothing at all
2 - Persuade people one-on-one
3 - Protest or apply public pressure for governments to find a diplomatic solution
4 - Protest or apply public pressure for governments to find a military solution (if issue in other country)/violently protest (if issue in same country)
5 - Support in spirit or financially a non-state militant group
6 - Join a non-state militant group
7 - Carry out militant actions to fight and die on my own, even if I had no group or support network.

<u>Islamic Teaching in Schools</u>
23) I **agree/disagree** that the Spanish government should offer Islamic teachings to Muslim children in public schools.

Please respond to one of the following scenarios based on whether you agree or disagree with the above issue.

Agree:
23.a.1) If instead of offering Islamic teaching in public schools the Spanish government offered large tax refunds to Muslim families so that they can provide Islamic tutoring for their children on their own (or spend the money any other way), or offered any other **purely financial** incentives, would you find it acceptable?

Yes Maybe No

23.b.1) If the majority of Muslim people in Spain decided that they don't need the Spanish government to provide Islamic teaching in public schools, would that be okay?

Yes Maybe No

Disagree:
23.a.2) If the Spanish government offered Islamic teaching in public schools because this would lead to more jobs for teachers and textbook companies and increase revenue to government through taxes, or offered any other **purely financial** incentives, would you find it acceptable?

Yes Maybe No

23.b.2) If the majority of Spanish people and the government decided that they wanted to provide Islamic teaching in public schools, would that be okay?

Yes Maybe No

23.c) Please select all the actions that you would be willing to take to defend your position on this issue.

1 - Nothing at all
2 - Persuade people one-on-one
3 - Protest or apply public pressure for governments to find a diplomatic solution
4 - Protest or apply public pressure for governments to find a military solution (if issue in other country)/violently protest (if issue in same country)
5 - Support in spirit or financially a non-state militant group
6 - Join a non-state militant group
7 - Carry out militant actions to fight and die on my own, even if I had no group or support network.

Mosques Building
24) I **agree/disagree** that the Spanish government should allow the unrestricted building of mosques.

Please respond to one of the following scenarios based on whether you agree or disagree with the above issue.

Agree:
24.a.1) If instead of allowing the unrestricted building of mosques the Spanish government offered a monthly stipend to Muslims so they could spend that money on creating decent prayer spaces within their own homes (or spend it any other way), or offered any other **purely financial** incentives, would you find it acceptable?

Yes Maybe No

24.b.1) If the majority of Muslim people in Spain decided that they don't need the Spanish government to allow the unrestricted building of mosques, would that be okay?

Yes Maybe No

Disagree:
24.a.2) If the Spanish government allowed the unrestricted building of mosques because this would lead to increased construction and development in Spain which would be good for the economy, or offered any other **purely financial** incentives, would you find it acceptable?

Yes Maybe No

24.b.2) If the majority of Spanish people and the government decided that they wanted to allow the unrestricted building of mosques, would that be okay?

Yes Maybe No

24.c) Please select all the actions that you would be willing to take to defend your position on this issue.

1 - Nothing at all
2 - Persuade people one-on-one
3 - Protest or apply public pressure for governments to find a diplomatic solution
4 - Protest or apply public pressure for governments to find a military solution (if issue in other country)/violently protest (if issue in same country)
5 - Support in spirit or financially a non-state militant group
6 - Join a non-state militant group
7 - Carry out militant actions to fight and die on my own, even if I had no group or support network.

Sharia
25) I **agree/disagree** that the strictest form of Sharia should be applied in all Muslim lands.

Please respond to one of the following scenarios based on whether you agree or disagree with the above issue.

Agree:
25.a.1) If a Muslim country chose not apply the strictest form of Sharia in their land because not doing so meant a better economy, more foreign trade-relations, more corporations willing to invest in the country, more foreign aid, etc., or any other **purely financial** reasons, would you find it acceptable?

Yes Maybe No

25.b.1) If the people and government of a Muslim country did not want the strictest form of Sharia applied in their country, would that be acceptable?

Yes Maybe No

Disagree:
25.a.2) If a Muslim country chose apply the strictest form of Sharia in their land because doing so led to economic benefits such as increased foreign aid, more jobs due to corporate partnerships, profit from the trade-relations etc., with other Muslim countries and their companies, or any other **purely financial** reasons, would you find it acceptable?

Yes Maybe No

25.b.2) If the people and government of a Muslim country wanted the strictest form of Sharia applied in their country, would that be acceptable?

Yes Maybe No

25.c) Please select all the actions that you would be willing to take to defend your position on this issue.

1 - Nothing at all
2 - Persuade people one-on-one
3 - Protest or apply public pressure for governments to find a diplomatic solution
4 - Protest or apply public pressure for governments to find a military solution (if issue in other country)/violently protest (if issue in same country)
5 - Support in spirit or financially a non-state militant group
6 - Join a non-state militant group
7 - Carry out militant actions to fight and die on my own, even if I had no group or support network.

Caliphate

26) I **agree/disagree** that the boundaries of current Muslim countries should be dismantled and replaced with a single Caliphate.

Please respond to one of the following scenarios based on whether you agree or disagree with the above issue.

Agree:
26.a.1) If a Muslim country did not want to dismantle its borders and join a single Caliphate because it felt that it could have a better economy, more jobs, more financial success, etc. by remaining an independent nation, or any other **purely financial** reasons, would you find it acceptable?

Yes Maybe No

26.b.1) If the people and government of a Muslim country did not want to dismantle its borders and a join a single Caliphate, would that be acceptable?

Yes Maybe No

Disagree:
26.a.2) If a Muslim country did want to dismantle its borders and join a single Caliphate because it felt that it could have a better economy, more jobs, more financial success, etc. by joining the Caliphate, or any other **purely financial** reasons, would you find it acceptable?

Yes Maybe No

26.b.2) If the people and government of a Muslim country wanted to dismantle its borders and a join a single Caliphate, would that be acceptable?

Yes Maybe No

26.c) Please select all the actions that you would be willing to take to defend your position on this issue.

1 - Nothing at all
2 - Persuade people one-on-one
3 - Protest or apply public pressure for governments to find a diplomatic solution
4 - Protest or apply public pressure for governments to find a military solution (if issue in other country)/violently protest (if issue in same country)
5 - Support in spirit or financially a non-state militant group
6 - Join a non-state militant group
7 - Carry out militant actions to fight and die on my own, even if I had no group or support network.

Jihad
27) I **agree/disagree** that armed jihad should be waged against the enemies of Muslims?

27.1) Select all groups that are the enemies of Muslims: USA, Israel, India, Bashar Al-Assad, Taliban, Al Qaida, ISIS, France, UK, Spain.

Please respond to one of the following scenarios based on whether you agree or disagree with the above issue.

Agree:
27.a.1) If instead of waging armed jihad against the enemies of Muslims, Muslims expanded the financial capabilities of Muslim countries by engaging in entrepreneurship, investment, trading in Muslim lands, etc., or any other **purely financial** reasons, would you find it acceptable?

Yes Maybe No

27.b.1) If the people and governments of the Muslim mmah did not want armed jihad to be waged against its enemies, would that be acceptable?

Yes Maybe No

Disagree:
27.a.2) If waging armed jihad against the enemies of Muslims led to greater economic benefits for all Muslims such improved economies due to the war effort, which means more companies, more jobs, more tax revenue, etc., or any other **purely financial** reasons, would you find it acceptable?

Yes Maybe No

27.b.2) If the people and governments of the Muslim ummah wanted to wage armed jihad against its enemies, would that be acceptable?

Yes Maybe No

27.c) Please select all the actions that you would be willing to take to defend your position on this issue.

1 - Nothing at all
2 - Persuade people one-on-one
3 - Protest or apply public pressure for governments to find a diplomatic solution
4 - Protest or apply public pressure for governments to find a military solution (if issue in other country)/violently protest (if issue in same country)
5 - Support in spirit or financially a non-state militant group
6 - Join a non-state militant group
7 - Carry out militant actions to fight and die on my own, even if I had no group or support network.

<u>Belief in Allah and Prophet Mohamed (SAW)</u>
28) I **agree/disagree** that everyone must profess that Allah is the one and only true God and Mohamed (SAW) is his messenger.

Please respond to one of the following scenarios based on whether you agree or disagree with the above issue.

Agree:
28.a.1) If in exchange for not having to profess belief in Allah and Prophet Mohamed (SAW) non-believers donated money to Muslim organizations such as Mosques, charities, community centers, etc., or offered any other **purely financial** incentives, would you find it acceptable?

Yes Maybe No

28.b.1) If some non-Muslims did not want to profess belief in Allah and Mohamed (SAW) as his messenger and the countries that they lived in were okay with this, would that be acceptable?

Yes Maybe No

Disagree:
28.a.2) If a previous non-believer who is a storeowner in a Muslim country did profess belief in Allah and Prophet Mohamed (SAW) because doing so led to many more Muslim customers for his store, or did so for **purely financial** reasons, would you find it acceptable?

Yes Maybe No

28.b.2) If some non-Muslims wanted to profess belief in Allah and Mohamed (SAW) as his messenger and the countries that they lived in were okay with this, would that be acceptable?

Yes Maybe No

28.c) Please select all the actions that you would be willing to take to defend your position on this issue.

1 - Nothing at all
2 - Persuade people one-on-one
3 - Protest or apply public pressure for governments to find a diplomatic solution
4 - Protest or apply public pressure for governments to find a military solution (if issue in other country)/violently protest (if issue in same country)
5 - Support in spirit or financially a non-state militant group
6 - Join a non-state militant group
7 - Carry out militant actions to fight and die on my own, even if I had no group or support network.

Caricatures of Prophet Mohamed
29) I **agree/disagree** that prophet Mohammed (SAW) must never be caricatured.

Please respond to one of the following scenarios based on whether you agree or disagree with the above issue.

Agree:
29.a.1) If a newspaper reprinted the caricatures of the prophet Mohamed (SAW) because they knew that if they didn't they would lose all financial support from their investors and advertisers and have to downsize their business and layoff some employees, or any other **purely financial** reasons, would you find it acceptable?

Yes Maybe No

29.b.1) If some magazine or newspaper wanted to caricature the prophet Mohamed (SAW) and the countries to which that magazine belonged were okay with this, would that be acceptable?

Yes Maybe No

Disagree:
29.b.2) If a newspaper did not reprint the caricatures of the prophet Mohamed (SAW) because they knew that if they did they would lose all financial support from their Muslim investors and advertisers and have to downsize their business and layoff some employees, or any other **purely financial** reasons, would you find it acceptable?

Yes Maybe No

29.b.2) If some magazine or newspaper did not want to caricature the prophet Mohamed (SAW) and the countries to which that magazine belonged were okay with this, would that be acceptable?

Yes Maybe No

29.c) Please select all the actions that you would be willing to take to defend your position on this issue.

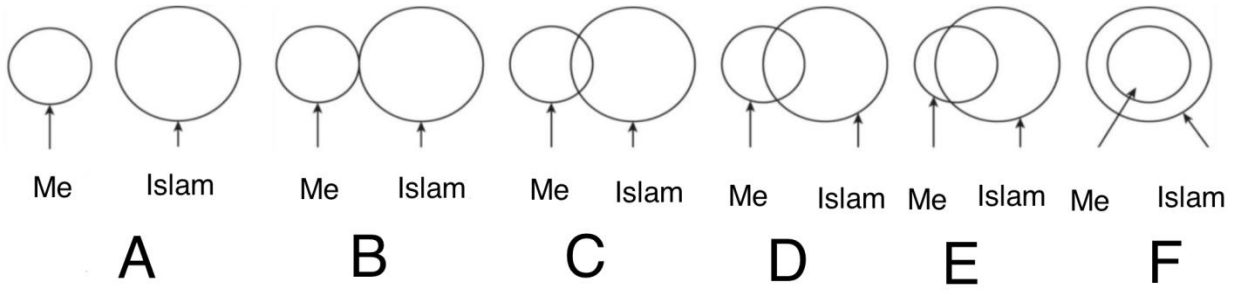
1 - Nothing at all
2 - Persuade people one-on-one
3 - Protest or apply public pressure for governments to find a diplomatic solution
4 - Protest or apply public pressure for governments to find a military solution (if issue in other country)/violently protest (if issue in same country)
5 - Support in spirit or financially a non-state militant group
6 - Join a non-state militant group
7 - Carry out militant actions to fight and die on my own, even if I had no group or support network.

30. Please rank the values below in order of importance. This ranking is the order in which you would choose the value if these values ever came into competition: The first value is the one you would select above all other values if you had to select only one to succeed, and so on.
Values: Kashmir, Gay marriage, US military, Drone strikes, Palestinian RoR, Palestinian statehood, Halal food, Islamic teaching, Mosque building, Sharia, Jihad, Allah, Caricatures

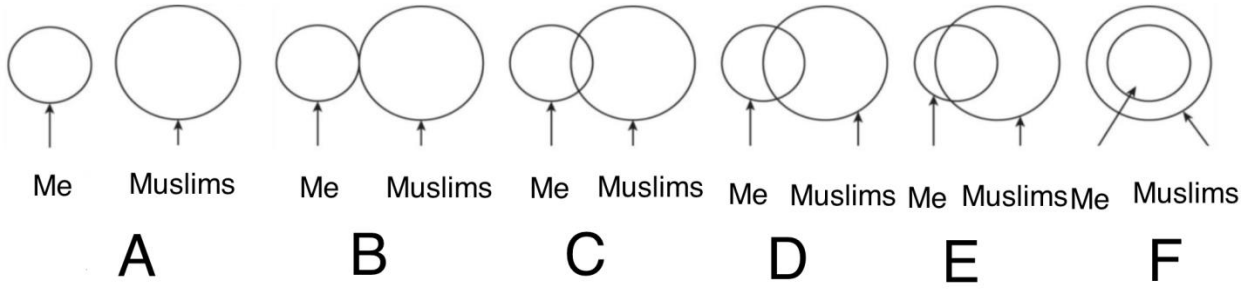
1. _____
2. _____
3. _____
4. _____
5 _____
6. _____
7. _____
8. _____
9. _____
10. _____

31. Here are two circles. The big circle represents a group the small circle represents you. Pick one these 5 pairs of small and big circles that best reflects your relationship with this group.

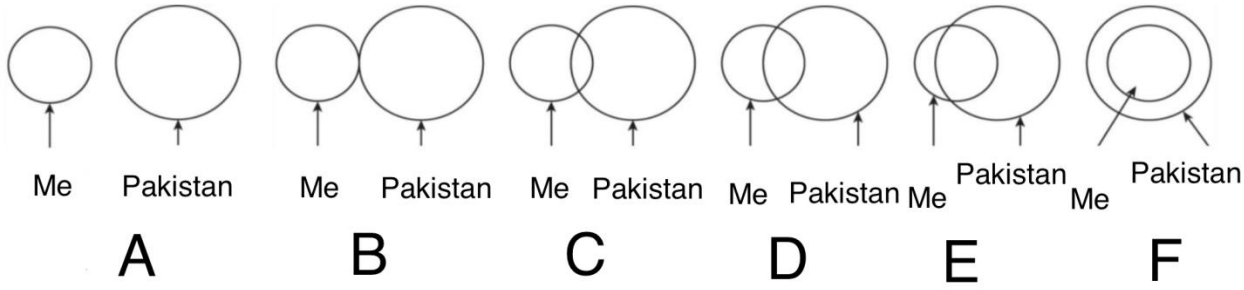
Islam



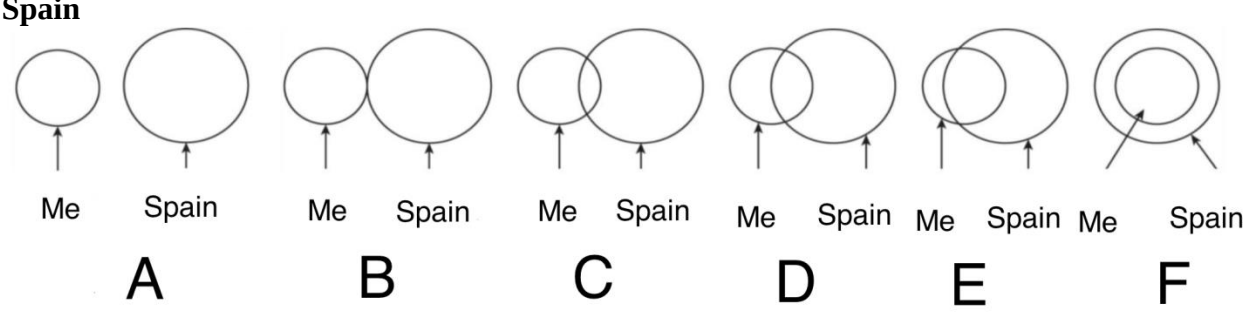
Muslims



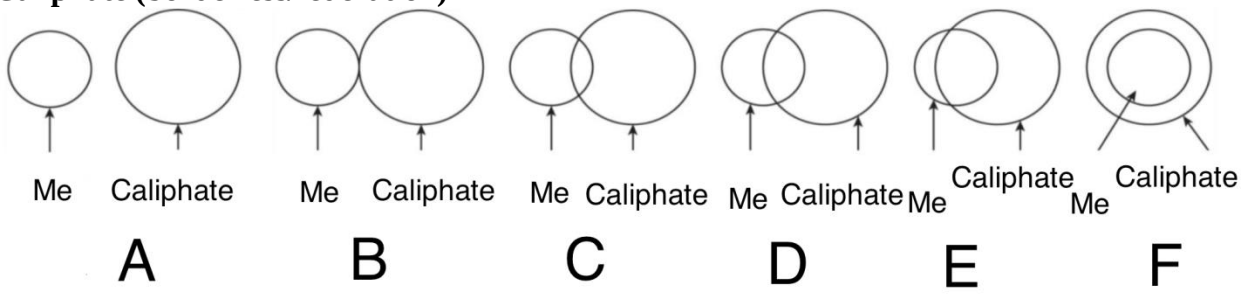
Pakistan



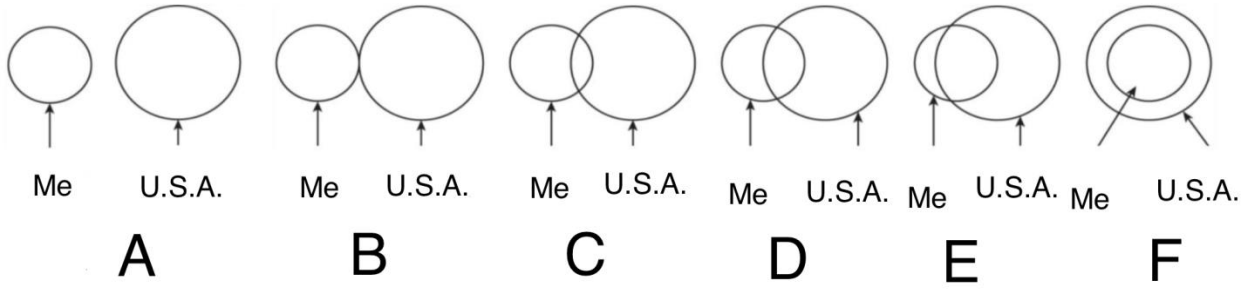
Spain



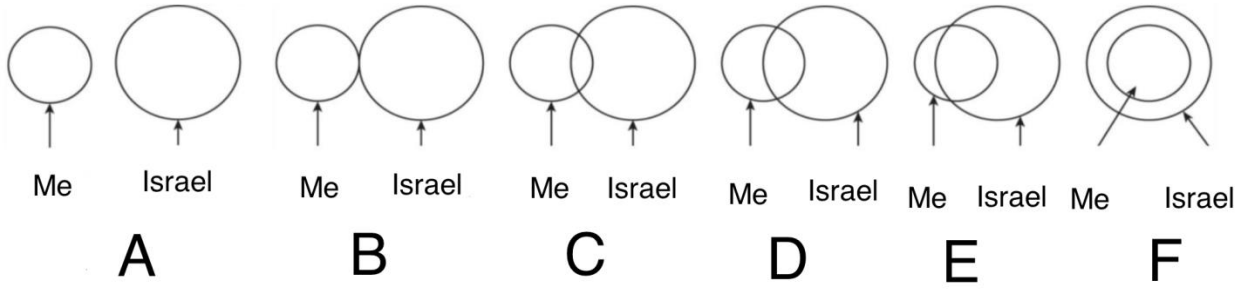
Caliphate (borderless/federation)



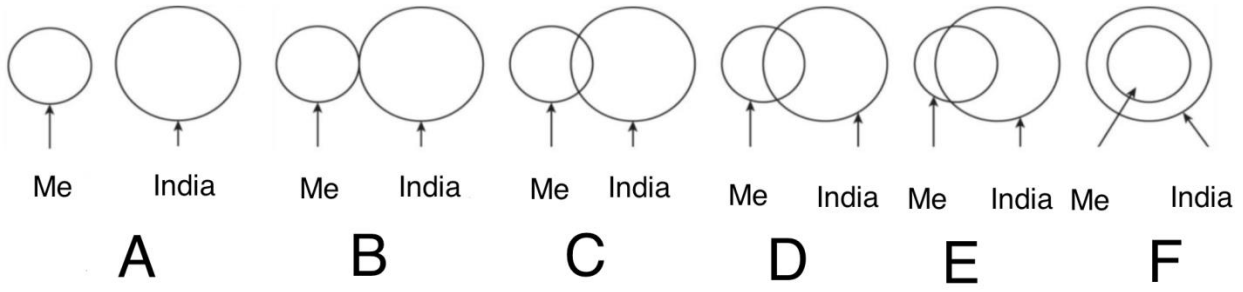
U.S.A.



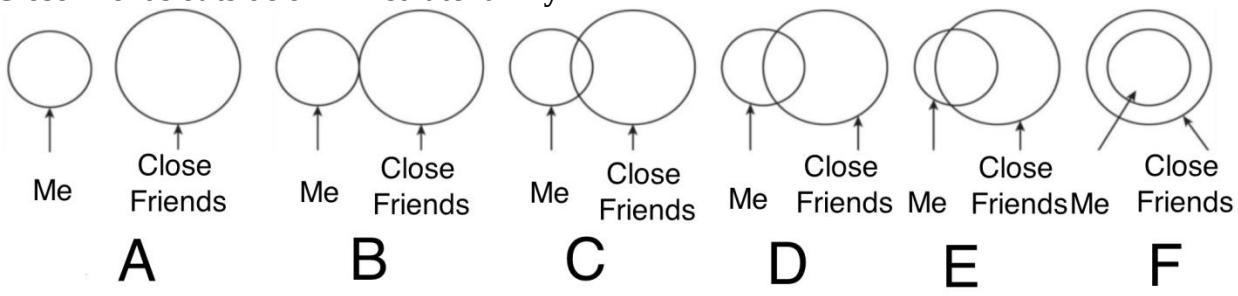
Israel



India



Close friends outside of immediate family



32. Please rank the following groups in order of importance to you: Islam, Muslims, Pakistan, Spain, Caliphate, USA, Israel, India, Close friends outside of immediate family.

1._____

2. _____

3._____

4._____

5._____

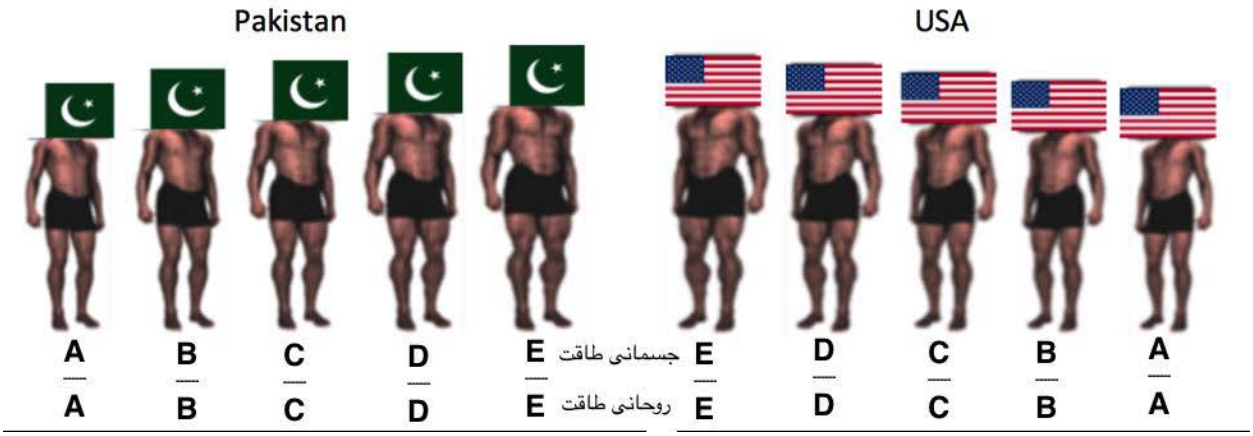
6._____

7._____

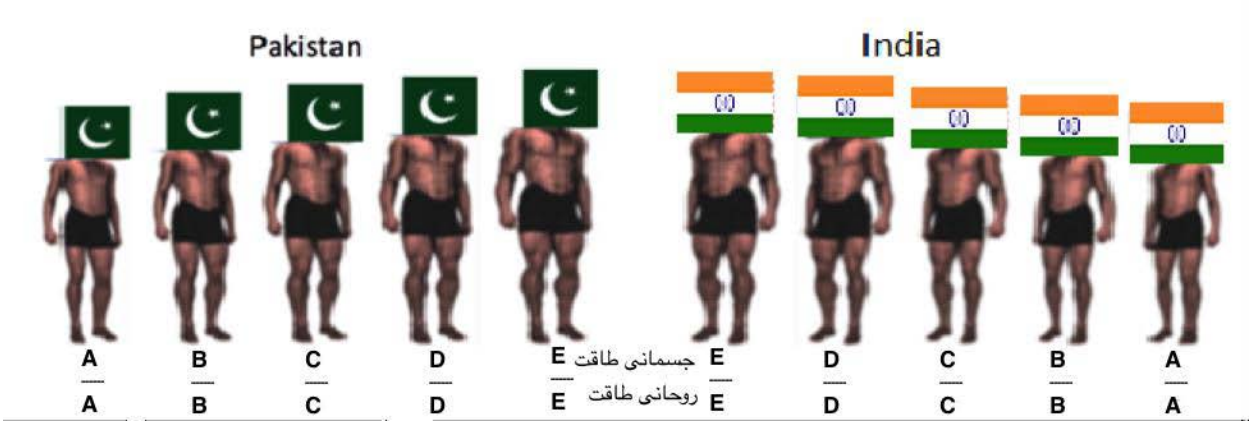
8._____

9._____

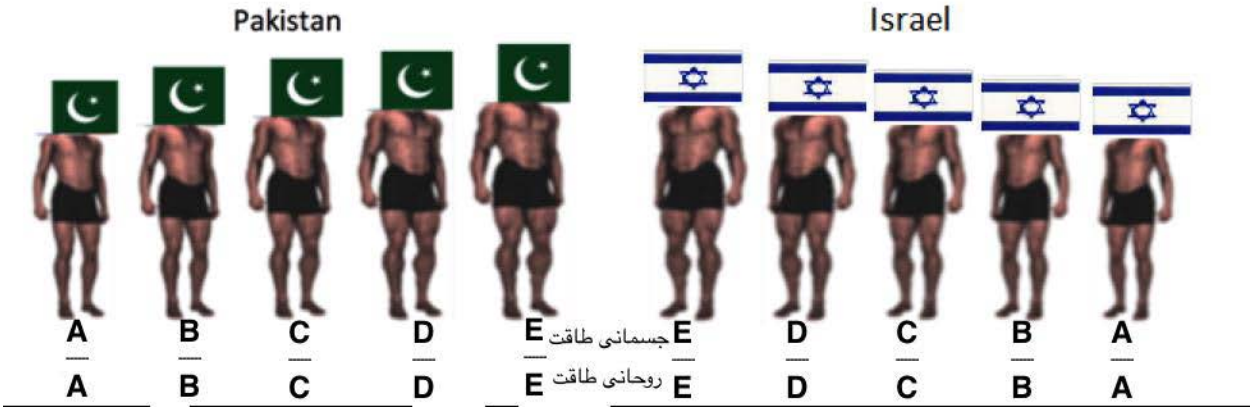
33. In the below diagram are two human bodies that represent the strength of two groups: **Pakistan** and the **USA**. Please select the figure that best represents the strength of each group. Please give us one rating of physical strength and one rating of spiritual strength for each group.



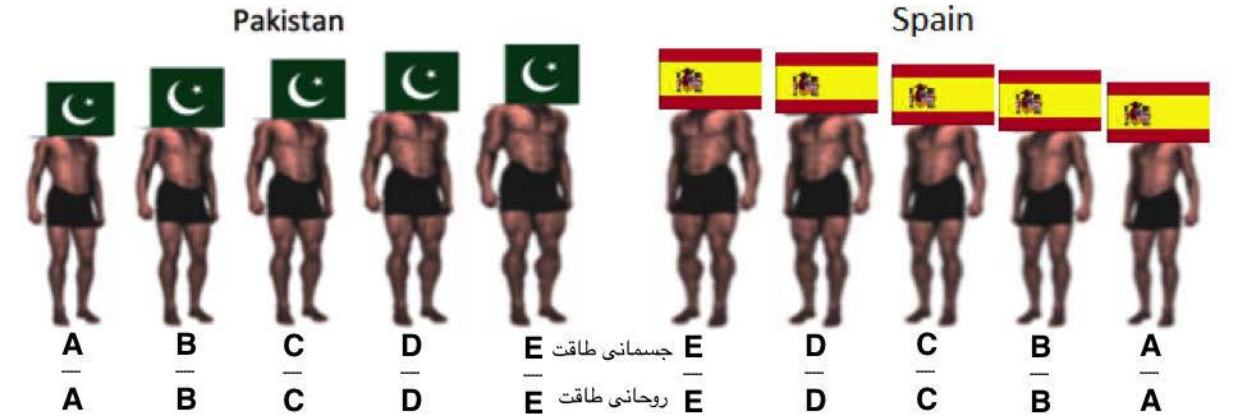
34. In the below diagram are two human bodies that represent the strength of two groups: **Pakistan** and **India**. Please select the figure that best represents the strength of each.. Please give us one rating of physical strength and one rating of spiritual strength for each group.



35. In the below diagram are two human bodies that represent the strength of two groups: **Pakistan** and **Israel**. Please select the figure that best represents the strength of each group. Please give us one rating of physical strength and one rating of spiritual strength for each group.



36. In the below diagram are two human bodies that represent the strength of two groups: **Pakistan** and **Spain**. Please select the figure that best represents the strength of each group. Please give us one rating of physical strength and one rating of spiritual strength for each group.



37. **Donation**
If we were to donate up to €10 on behalf of each participant, which charity or organization would you want us to give your donation to from the previous list of groups or parties?

_____

38. How much of your free time do you spend keeping up with world politics (reading blogs, newspapers, watching the news, etc.), especially the news concerning Muslims around the world?

1 All my free time
2
3
4
5
6
7Not at All

39) Which of the following statements do you agree with more:

a) Religious texts such as the Quran and Hadith must be viewed in terms of their historical context and be studied to find their meaning in a modern context

b) Religious texts such as the Quran and Hadith must be taken "as is" and all rules must be followed without adjustment to modern contexts.

40) Which country do you see as the most ideal Muslim country; the country you would like to see Pakistan become more like?

_____