

hypotheses, but there remains a need for wider independent replications to validate existing findings.

The pathways to fusion and self-sacrifice proposed here could turn out to be mistaken in some of their details without being completely wrong. What would be fatal for the theory is if it turned out that convictions of shared essence failed to predict high fusion scores, or if fusion (plus out-group threat) were shown to be a poor predictor of actual (as opposed to declared) willingness to fight and die for the group. These two claims are so central to the conceptual framework that, if shown to be false, the entire edifice would collapse. Somewhat less disastrous for the theory, but still a setback, would be a significant reduction in its explanatory provenance. For example, the theory may eventually prove to be applicable only to some armed groups but not all, and perhaps most crucially not to suicide terrorists. Although there would seem to be many similarities between the self-sacrificial acts of armed militia (whose fusion levels with numerous target groups have been measured) and those of suicide bombers (whose fusion levels are unknown), these similarities may turn out to be more apparent than real. If, as some have argued (see sect. 2), most suicide terrorists are motivated by pathology (e.g., suicidal depression) rather than the desire to act in the interests of a group, that would be a serious problem for the theory as articulated in this paper. Decisive evidence on this question may require more extensive research among would-be suicide terrorists and those who have attempted unsuccessfully to carry out such attacks (the previously acknowledged difficulties of conducting such studies notwithstanding).

The theory presented here also raises many new, empirically tractable questions, for example, concerning the relationship between local and extended fusion. Future research should investigate whether perceptions of shared essence are stronger if they are based on direct observation rather than on the testimony of others. Would remembering who else was there alongside you in a decisive battle or a traumatic rite of passage or perceiving shared phenotypic traits in a sibling provide more compelling evidence of shared experience or shared biology than merely displaying the same kind of medals or reciting myths of shared ancestry? Relational ties to a local group often incorporate episodic memories for self-defining events, which other group members indelibly inhabit. By contrast, categorical ties to an extended group are based largely on “knowing that” certain identity markers serve as indirect testimony to shared experience. Indirect evidence of shared experience may not be capable of motivating acts of self-sacrifice to the same extent as bonds forged through episodic memories of shared ordeals within a band of brothers.

Research into the causes of extreme pro-group action is not merely of scientific interest; there is potential also to use the findings in practical ways. For example, deradicalising Islamist militants might be reframed as a process of *defusing* extremists. Given that we now have a well-substantiated account of the causal pathways to fusion, together with evidence that priming the mediating variables in this pathway increases fusion (Whitehouse et al. 2017), it may be possible to reduce the effects of mediating variables so as to obstruct or reverse the fusion process. This has yet to be demonstrated in practice, but the general approach is well motivated theoretically. Such an approach should not be confused with the notion of “deprogramming” because the goal would not be to alter people’s beliefs or goals against their will. Indeed, the aim would not be to challenge the validity of ideologies or doctrines at all, but only to facilitate a process of reflection on past experiences and their relevance to group alignments. The process would need to engage the

wider participation not only of extremists but also of members of their social networks and surrounding communities (such as parents, schoolteachers, and religious leaders), although the ethics of any interventions would require careful scrutiny and monitoring.

Yet another potential application of this new framework would be neither to create nor to obstruct group alignments but to harness existing ones. There are a number of potentially desirable ways in which this could be done, not least in rebuilding societies devastated by conflicts or natural disasters. For example, during the uprising of 2011, many Libyans fought passionately and at huge cost to clear the way for a prosperous future under a more consensual system of governance. The social cohesion needed to build that vision was available in abundance at the end of the revolution, but there was a failure to harness it for the public good both on the part of the international community and on the part of Libyan leaders vying for power at the time. The same pattern repeats itself endlessly in other conflicts around the world. Only by better understanding the underlying causes of pro-group commitment can we benefit from its potential for building trust and cooperation while limiting its capacity to stoke intergroup conflict.

Acknowledgments. Thanks to Michael Buhrmester, Sergey Gavrilets, Rohan Kapitany, Christopher Kavanagh, Jonathan Lanman, and the BBS reviewers for comments and suggestions. This work was supported by a Large Grant from the UK’s Economic and Social Research Council (REF RES-060-25-467 0085), an Advanced Grant from the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation Programme (Grant Agreement 694986), and an award from the Templeton World Charity Foundation entitled “Cognitive and Cultural Foundations of Religion and Morality” (TWCF0164).

Open Peer Commentary

What motivates devoted actors to extreme sacrifice, identity fusion, or sacred values?

Scott Atran^{a,b} and Ángel Gómez^{c,d} 

^aCentre National de la Recherche Scientifique, Paris, France; ^bChanging Character of War Centre, University of Oxford, Oxford, United Kingdom; ^cUniversidad Nacional de Educación a Distancia, Madrid, Spain and ^dArtis International, Scottsdale, AZ 85824.

scott.atran@pmb.ox.ac.uk

<http://www.hmc.ox.ac.uk/people/professor-scott-atran/>

agomez@psi.uned.es

<https://angelgomezjimenez.webnode.es/>

doi:10.1017/S0140525X18001565, e193

Abstract

Why do some individuals willingly make extreme sacrifices for their group? Whitehouse argues that such willingness stems from a visceral feeling of oneness with the group – identity fusion – that emerges from intense, shared dysphoric experiences or from perceived close kinship with others. Although Whitehouse’s argument makes a valuable contribution to understanding extreme sacrifice, factors independent of identity fusion, such as devotion to sacred values, can predict self-sacrifice.

Ever since Darwin (1871), scientists have puzzled over why some people, such as heroes and martyrs, willingly self-sacrifice, even when facing overwhelming odds and apparent defeat. The global advent of suicide attacks has transformed the issue into a paramount policy challenge for governments and their publics. Whitehouse's article is informative and timely, focusing its explanation of violent extremism on an interrelated complex of cognitive and emotional means for binding groups (perceptions of shared essence, actual and imagined kinship, shared episodic memories, and intense emotional experiences), while offering general understanding of self-sacrifice applicable to many cultural contexts and times. However, Whitehouse risks overstating his case by claiming that identity fusion is the primary, if not unique, driver of extreme sacrifice.

In the last decade, experiments performed on five continents have shown identity fusion is a reliable predictor of willingness to fight, kill, and die for one's group. Identity fusion theory originated with William Swann and Ángel Gómez in 2005. It was initially conceived to help explain the September 11, 2001 attacks and March 11, 2004 Madrid train bombings (Europe's worst terrorist attack to date). It was then empirically validated in several publications (Gómez et al. 2011a; Swann et al. 2009). Whitehouse subsequently joined the effort (Swann et al. 2012), applying the theory with colleagues to an impressive set of field settings from initiation rites in New Guinea to the Libyan insurgency against Gaddafi (Whitehouse et al. 2014b).

The target article is compelling when extending fusion theory to explain the group-binding functions of intense, dysphoric experiences in painful rituals or other emotional life-shaping experiences (e.g., frontline combat). Whitehouse convincingly relates such experiences to kin psychology: attitudes and feelings associated with immediate familial ties, which can be extended to larger groups – from tribes to transnational movements – via participation in intensely emotional rituals or attention to symbols that evoke shared intense experiences. Previous fusion research supports the connection between these mechanisms and fusion. For example, individuals diagnosed with gender dysphoria (i.e., transsexuals), when fused with their preferred gender, are willing to suffer painful experiences (e.g., major surgery) to belong to their desired sex group (Swann et al. 2015). Other studies also show that fusion promotes self-sacrifice, including dying for a group, by fostering perception of familial ties (Swann et al. 2014a).

Less compelling is Whitehouse's argument that identity fusion is generally the principal determinant of willingness to self-sacrifice. Other anthropological and psychological research indicates that commitment to so-called sacred values can motivate extreme and costly behaviors (Baron & Spranca 1997; Graham & Haidt 2013; Rappaport 1971; Tetlock 2003). Whitehouse dubiously acknowledges sacred values by assimilating them to identity fusion. Thus, "extreme beliefs [may] become so closely linked to the group that they take on an aura of sacredness"; however, "what connects those values to acts of self-sacrifice may well be fusion with the group rather than commitment to any kind explicit belief system" (sect. 2, para. 6).

Yet, among Itza' Maya in lowland Guatemala, we find strong commitment to spiritual values that summarize millennial experience – but no significant contemporary group bonding, ritualized or otherwise – driving very costly rainforest management (Atran

et al. 2002). Studies in Western Europe, North Africa, and the Middle East reveal sacred values and identity fusion to be uncorrelated, independent predictors of willingness to engage in, and suffer, extreme violence. When individuals perceive a threat both to their fused group and to sacred values, identity fusion and sacred values interact, leading to greater willingness to sacrifice than for either factor alone (Atran et al. 2014; Sheikh et al. 2016). Sometimes identity fusion takes precedence over sacred values (Gómez et al. 2016a). In other circumstances, sacred values prove more important. For example, in our study of frontline combatants in Iraq (Kurdish PKK and Peshmerga, Sunni Arab militia, Iraqi Army, captured Islamic State fighters), those most willing to make costly sacrifices (as verbally expressed and in terms of actually being wounded and voluntarily returning to fight) were ready to forsake their fused group, whether their genetic family or any other group with which they were fused, rather than their sacred values. This finding was replicated among subjects most willing to make costly sacrifices in a sample of more than 6,000 Western Europeans (Gómez et al. 2017) and with young men just emerging from Islamic State rule in the Mosul area of Iraq (Atran et al. 2018).

Whitehouse questions these findings, arguing that "measures of sacred values ... are related to similar measures of willingness to sacrifice for sacred values" (sect. 2, para. 6). Our sacred value measures chiefly concern unwillingness to trade the value against material gain or loss (Ginges et al. 2011), although in some studies additional indicators of sacredness include insensitivity to discounting, immunity to peer pressure, and blindness to exit strategies (Sheikh et al. 2013). But in our frontline studies, for example, we see no support for Whitehouse's intimation that refusing material incentives for assessing sacred values, such as Sharia law, is conflated with outcome measures of costly commitments such as "dying, letting one's family suffer, undertaking a suicide attack, torturing women and children" (Gómez et al. 2017, p. 678).

Whitehouse surmises: "willingness to fight and die is not motivated by doctrines and ideologies, religious or otherwise, but by a particularly intense love of the group" (sect. 2, para. 7). Previous research suggests that even for some suicide attacks in the name of religion or for a political goal, group dynamics can be more important than confessional or ideological affiliation (Atran 2010; Sageman 2004). But in other circumstances, devotion to sacred values may be primary (Atran et al. 2018; Gómez et al. 2017) or be important even without any longstanding relationship to religious or ideological doctrine (e.g., right to nuclear capability among some Iranians [Dehghani et al. 2010]). Whitehouse (2000) distinguishes ideologies and doctrines from the imagistic and emotion-laden aspects of ritual and dysphoric experiences that, by and large, distinguish spiritual life in small-scale societies (e.g., pre-state cultures, contemporary New Guinea tribes) from the "doctrinal" religions and political ideologies of large-scale societies (e.g., empires, nations). Sacred values, though, appear to have privileged connections to emotions and can be as imagistic and intensely felt (Atran & Ginges 2012; Durkheim 1912; Ginges et al. 2007; Gómez et al. 2017), as they can be part of religious or ideological doctrine.

A general theory of extreme self-sacrifice should consider, at a minimum, that people can make extreme sacrifices for a group, but also, or even independently, for a cherished cause.