<R-Header>Author's Response

<RT>Four things we need to know about extreme self-sacrifice

<RA>Harvey Whitehouse

<RAA> School of Anthropology and Museum Ethnography, University of Oxford, Oxford, United Kingdom

harvey.whitehouse@anthro.ox.ac.uk

<R-AB>Abstract: A comprehensive explanation of extreme self-sacrifice would ideally clarify not only the proximate mechanisms leading to this behaviour, but also its developmental origins, its functions (if any), and its history. The theoretical framework set out in my target article has something to say about all these dimensions, and many of the criticisms raised in the commentaries can therefore be addressed under those four main headings. My response also discusses a set of proposals for further extending the framework. Finally, it concludes, by distilling from the discussion, a host of novel questions requiring further investigation.

<R-Text Begins>According to the Nobel prizewinning ethologist Nico Tinbergen, a comprehensive evolutionary account of any phenomenon must address four fundamental questions (Tinbergen 1963): (1) What are immediate proximate causes of the phenomenon? (2) How is it assembled over the course of development? (3) What is its evolved function, if

any? (4) What is its evolutionary history? These questions can be asked, not only of biologically evolved characteristics (e.g. organs and instincts), but also of culturally evolved features (e.g. institutions and artefacts). All four questions have helped to shape the theory of extreme self-sacrifice set out in my target article. Accordingly, many of the critical commentaries respond to, and can be grouped around, each of Tinbergen's four questions in turn (see Table 1). Towards the end of this response, I discuss separately, commentaries seeking to extend, rather than to critique, the pathways-to-fusion framework (see Table 2, section 5). I conclude with a short summary of the implications of this wide-ranging discussion for future research on the causes of extreme self-sacrifice.

<RA> R1. **On mechanism**

<RB> R1.1. ***Shared values and beliefs can motivate extreme self-sacrifice***

Numerous commentators observed that shared values and beliefs might play an important role in motivating extreme self-sacrifice, either independently of fusion or in tandem with it. Boudry, for example, points out that the mass suicides associated with the Heaven's Gate cult or the People's Temple in Jonestown make little sense except in the light of a belief in a better afterlife. But before we can conclude that extreme behaviors in cults are generally a consequence of extreme beliefs, we should also consider coercive leadership, extreme peer pressure, and potentially many other factors. Would any one of the cult followers in the groups mentioned above have gone ahead with their personal suicide mission if all other cultists had decided against it? One of the most robust discoveries of experimental psychology, taken in the round, is that belief alone is a poor predictor of behavior and often constitutes little more than post hoc rationalization. Accordingly, the teachings of cult leaders

may be powerless to motivate behavior unless somehow tapping into deeper motivational systems, for example, rooted in kin psychology or us-them coalitional thinking. In short, although there could be multiple factors driving extreme behavior, of which fusion-plus-threat may only be one, we cannot simply assume that extreme beliefs exert a privileged influence independently of other factors. We need evidence to support this.

To this end, Atran & Gómez have made the case that certain kinds of values (specifically those that have become somehow sacralized), can lead to extreme self-sacrifice and, more to the point in the present context, can do so independently of the effects of fusion. They may well be right about this, but there are reasons to be cautious about the sacred values explanation for self-sacrifice, at least for now. First, the theory of sacred values is currently underspecified, in that we do not know how values become sacralized in the first place, such that people might be willing to fight and die for them. By contrast, we have an elaborated conceptual framework specifying the causal pathways to fusion (summarized in Fig. 1 of the target article), and a growing body of empirical data to support the causal arrows in that model. Second, despite some preliminary data suggesting that fusion and sacred values have distinct psychological effects and behavioral outcomes, evidence that sacred values motivate extreme self-sacrifice in the absence of fusion, is lacking or inconclusive. This is not to say that there will never be a theory of what makes certain values sacred and how they can motivate self-sacrifice independently of group alignments. We should continue to explore these issues through rigorous empirical research, building on existing foundations (Atran et al. 2014; Gomez et al. 2017).

The role of values is emphasized also by Swann & Jetten. But rather than arguing that values operate independently of fusion, they suggest that shared values can lead to fusion in

much the same way as shared experiences and shared biology. In the context of the general conceptual framework I am advancing, however, sharing cultural beliefs and practices, as well as the group's values, should lead to identification, rather than fusion, as part of what I have called the 'doctrinal mode of religiosity' (Whitehouse 2004). Unlike transformative experiences in the imagistic mode, which are both intensely personal as well as shared with others, doctrinal beliefs and practices are acquired through social learning, and therefore, defining for the group but not for the essential autobiographical self. Nevertheless, one could imagine how a group's beliefs, practices, and values could become personalized, for example, through participation in commemorative rituals that evoke memories of a shared event (such as a personally experienced civil war), or nationalistic rallies that evoke notions of shared ancestry (such as a common motherland). Therefore, although the distinction between imagistic and doctrinal modes suggests that fusion and identification are rooted in quite different social and psychological processes, including distinct systems of cultural transmission and memory (Whitehouse 1992, 2004; Whitehouse & Lanman 2014), the overall framework I am proposing also opens up a range of other possible ways in which values and beliefs can become entangled in both forms of group alignment.

The idea that a group's beliefs can become linked with group identity, including group defining properties or events, is also captured by the notion of "conceptual ties," suggested by Lane et al. They argue, that while fusion with a local clique based on relational ties may explain some instances of radicalization, often individuals acquire extreme beliefs via the internet. Lane et al suggest that the latter may involve a fusion between the personal self and a set of concepts, as espoused online by terrorist organizations. If so, that might reveal yet another distinct pathway to fusion and self-sacrifice. Nevertheless, as the authors

of this commentary acknowledge, a lot more empirical work would be needed before such a conclusion could be drawn.

<RB> ***R1.1. Communal sharing rather than fusion explains self-sacrifice***

According to Thomsen & Fiske, there is a longer-established explanation for self-sacrifice, known as "communal sharing," that obviates the need for fusion theory. Communal sharing proposes a pathway to shared essence based on physical intimacy, cohabitation, commensality, and shared experience. Although Thomsen & Fiske say that communal sharing rarely involves pain, fear, or other dysphoric experiences, they mention, as examples of communal experiences, that give rise to shared essence such practices as "blood brotherhood rituals," circumcision, clitoridectomy, and giving birth. We have collected data showing that at least some of these kinds of life experiences do indeed have dysphoric elements that can contribute to processes of fusion with others who have also undergone them (Whitehouse & Lanman 2014; Whitehouse et al. 2017; Kavanagh et al. 2018). Those forms of communal sharing mentioned by Thomsen and Fiske that do not fit with our imagistic-pathway-to-fusion model, would nevertheless seem to fit well with our shared biology pathway, insofar as they emphasize bonds of kinship via nursing, feeding, commensalism, and so on (Vázquez et al. 2017). The main exception to this is social synchrony, which the pathways-to-fusion framework also encompasses (Jackson et al. 2018; Reddish et al. 2016), but in a model leading to state fusion rather than trait fusion (Swann et al. 2012). State fusion results from a temporarily elevated experience of shared essence, which, in the case of social synchrony, appears to be prompted by the illusion that one's agency and that of the group are combined (Reddish et al. in prep). When the experience of social synchrony ends, however, the effect goes away. Trait fusion, by contrast, is an

enduring form of group alignment rooted in long-lasting episodic memories (van Mulukum et al. in prep), protracted reflection (Jong et al. 2015), and phenotypic matching (Whitehouse & Lanman 2014). The theory of communal sharing lumps all these distinct elements together and, arguably, underspecifies the mechanisms by which they lead to self-sacrifice. Nevertheless, if theories of communal sharing and of multiple pathways to fusion were to make contrasting predictions, then it would be good to devise ways of adjudicating between them experimentally.

<RB> *R.1.1 Segregation not fusion increases the effects of out-group threat on extreme self-sacrifice*

Two of the commentaries suggest that segregation could moderate the relationship between out-group threat and extreme self-sacrifice. Kiper and Sosis cite research showing that reduced contact with an out-group is associated with increased risk of violent expressions of hostility. The idea is developed more fully by Melton and Motyl, who argue that when groups form enclaves, they adopt increasingly hostile attitudes toward out-groups, whereas, by contrast, when groups regularly encounter rival groups in peaceful settings they are less likely to develop extremist and violent attitudes toward out-groups. As argued above, however, it is important to distinguish hostile beliefs, norms, and attitudes toward out-groups, from actual willingness to fight and die for one's in-group. Melton and Motyl appear to regard these as more or less synonymous, but perhaps a more plausible hypothesis would be that the isolation of in-groups is more likely to lead to violent self-sacrifice when in-group members are highly fused and also threatened by an out-group. In other words, our respective frameworks may work better in combination, rather than as rival explanations. Future research should explore the possibility that enclave formation moderates the relationship

between fusion and out-group threat, on the one hand, and extreme self-sacrifice as outcome variable, on the other. Again, however, the question will be how much greater predictive power can be gained by adding additional moderators to the conceptual framework.

<RB> R1.1. Self-sacrifice is not always about in-groups

Several commentators observed that people sometimes make sacrifices for the sake of others who are not members of their in-groups. Crimston & Hornsey, for example, summarize research showing that people who score high on a "moral expansiveness scale" (measuring the range of entities one considers worthy of moral concern and care), were also more likely to express a willingness to donate organs or even give up their lives for members of out-groups, other animal species, and even to save trees (Crimston et al. 2016). Arguably, a limitation of the moral expansiveness research is that it only measures norms concerning the moral worth of various entities relative to self rather than group, and these norms may not predict behavior (i.e. people may not practice what they preach).

Even setting aside those concerns, it is not entirely clear how Crimston & Hornsey's observations relate to the arguments advanced in the target article. Previous research has shown that fusion can be extended not only to very large in-groups (such as nation, ethnic group, or world religion), but also to other species (Buhrmester et al. 2018) and even to supernatural beings (Buhrmester & Lanman in prep). Nevertheless, in all these cases, the fusion target may be describable as an in-group, from the fuser's perspective. For example, although social synchrony has been shown to generate fusion with out-groups (Reddish et al. 2016), at the moment when "state fusion" is measured, the out-group seems to be functioning

as an extended in-group (e.g. humanity at large), albeit temporarily. As such, fusion could conceivably motivate self-sacrifice toward groups usually regarded as rivals or enemies.

Relatedly, Elnakouri et al. observe that some people fight for the rights of out-groups (e.g. white Americans who campaigned for the abolition of slavery), even if this means risking life and limb. But again, it seems plausible that those who join the fight against social injustices are in fact strongly aligned with the downtrodden and oppressed, far from regarding them as an out-group. To the extent that white people in the past were willing to put their lives on the line to abolish slavery in America or apartheid in South Africa, they may indeed have been fused with oppressed groups despite differences of culture and ethnicity. Although we cannot directly measure fusion in historical populations no longer living, we do have evidence that fusion can cut across cultural and even species boundaries (Buhrmester et al. 2018).

<RB>R1.1. Fusion does not always lead to violent extremism

A few commentators mistook me to be arguing that fusion is invariably associated with violent extremism. Indeed, Kiper and Sosis went so far as to suggest that fusion and extreme self-sacrifice are essentially the same thing, and therefore judged the argument to be circular. The reality, however, is that fusion is a psychological construct, and extreme self-sacrifice is a behavior. One could logically occur in the absence of the other, and my argument was, in fact, that they do. For example, it was emphasized that fused individuals do not engage in extreme self-sacrifice unless a plausible threat is present, and also that there are other potential causes of self-sacrifice besides fusion. The same might be said in response to Xygalatas' observation that fused members of university sororities and fraternities generally

do not carry out suicide attacks. The argument is not that fusion always leads to suicide bombings, still less that fusion and self-sacrifice are the same thing. On the contrary, fused groups do not care very much about out-groups unless they present a danger. We know little about how members of fraternities or sororities would respond if they were persecuted or threatened with violence.

A related criticism, however, is that even if fusion does motivate violence or extreme self-sacrifice, it might not motivate both. Hansen argues that fighting and dying are conceptually distinct and perhaps even negatively related. It is indeed true that one can be willing to fight but not die, or to die but not fight. But the reason the two are conjoined in our conceptual framework is because a host of previous research has shown that fusion predicts willingness to fight and die (i.e. both together), as measured using a single well-validated "fight and die" scale (Swann et al. 2014). Hansen's hunch that fusion may better predict "die than fight," is intriguing, but we lack direct evidence of this using fusion measures, so this is something to explore in future research.

Several commentators pointed out that fusion can motivate peaceful as well as violent forms of pro-sociality. Swann and Jetten argue that much more could be said on that topic, as I too have argued elsewhere (Whitehouse 2013). Nevertheless, the present target article was primarily attempting to explain why some people will fight and die for their groups rather than to catalog all the "nice" outcomes of loving one's group. The same may be said in response to Olivola's observation that people make many kinds of sacrifices, not just violent ones, as well as Thomsen & Fiske's point that the sharing of traumatic experiences can lead to caring and compassion rather than to violence. Again, I agree, and this is precisely why I argued that when fusion is directed toward antisocial outcomes, as in the case of football

hooliganism, it might be possible to channel it in more positive directions, for example, in support of charitable causes. Olivola insists that self-sacrifice for charities is not motivated by fusion, but surely that is an open question empirically? Several studies have indeed shown that fusion can be a potent motivator of charitable giving, including donations of blood and money (Buhrmester et al. 2013), not only to assist strangers but even to help protect other species (Buhrmester et al. 2018). So, the point is well taken that a broader discussion of the role of fusion in motivating costly prosocial action should encompass research into these forms of peaceful self-sacrifice.

<RB> R1.1. Shared trauma can trigger group bonding instantaneously

According to the conceptual framework I have presented, shared experiences (perhaps shared suffering in particular), lead to fusion gradually over time, via processes of remembering and reflecting on the experience for a long period afterwards, with the result that the event and its significance form an indelible part of one's personal identity as well as one's group identity. Several commentators argued, by contrast, that the link between trauma and social cohesion may be much more immediate. Consider, for instance, the idea that when the group comes under attack, its members will spontaneously band together, realizing that their chances of survival will increase if they stick together. Lankford argues that this, rather than identity fusion, explains social bonding in response to shared dysphoric experiences. Lankford's argument fits quite well with the idea, discussed in the target article, that identification motivates ingroup bias and out-group derogation, albeit, not at the cost of individual survival. From a "strength in numbers" perspective, one should cling to the group as long as it remains strong, but defect if it weakens. Therefore, when two individuals are isolated from the herd, it makes evolutionary sense for the stronger individual to hamper the weaker one if that will

improve its own chances of escape. Efforts to disguise these inconvenient truths might well lead, as Lankford suggests, to overinflated claims of loyalty and willingness to self-sacrifice. But even if such claims sometimes amount to empty hyperbole and bravado that does not mean that they are never sincere.

The target article summarized numerous studies showing that fused individuals really are willing to lay down their lives for their groups. This willingness, however, is not triggered instantly in response to the appearance of a threat, as predicted by the strength in numbers hypothesis. Fusion comes about gradually, as a consequence of reflection on transformative experiences shared with other group members. (Jong et al. 2015; Buhrmester et al. 2018; van Mulukom et al. in prep; Muzzulini et al. in prep). As a result of these processes, fused individuals show a strong tendency to jeopardize not only their own welfare but also that of other group members, in order to act in the interests of the group as a whole. A good example of this would be whistleblowers (Buhrmester, 2013), who, if acting purely in the interests of self-preservation, would be far better off looking the other way, just like many others around them. The strength in numbers argument cannot account for this.

Although Swann and Jetten agree that sharing dysphoric experiences can lead to fusion, like Lankford, they suggest that it does so instantaneously as a kind of triggering effect. And from that perspective, they observe that dysphoric experiences might just as easily lead to antisocial behavior, such as random violence and looting. True, there may be many ways in which people respond to disasters in the heat of the moment but the shared experience pathway to fusion is about the long-term effects of such events on the formation of the personal self and on group identity. As already noted, recent longitudinal studies provide compelling evidence that the process of fusing with a group in this way is gradual

rather than instantaneous, and this is key to explaining the enduring nature of trait fusion. Far from being a knee-jerk reaction to one's immediate environment, trait fusion is a result of fundamental and durable changes to one's identity that take significant time to assemble (see above).

While the main focus of the target article was on trait fusion, there is also burgeoning literature on so-called state fusion, a temporary sense of oneness with others following a priming event but which fades soon after the event has passed. Such events are typically positive, even euphoric, rather than traumatic. The commentary by Xygalatas focuses on examples of such events in the context of spectator sports, in which huge crowds engage in singing, chanting, and other collective rituals. Xygalatas argues that the state fusion generated by such events may be neither local nor extended. This is an interesting point. A recent study of the effects of social synchrony (which of course abounds in football crowds), found that moving in time with others increases state fusion, not only toward co-participants but also toward anonymous members of an extended in-group (Reddish et al. 2016). More strikingly, those assigned to the synchronous condition in this study experienced elevated state fusion even toward members of an out-group. This suggests that certain kinds of collective experiences can increase fusion with others in general. It is possible that military rituals elevating state fusion prior to combat have long helped motivate warriors to fight and die on the battlefield, not because they foment hatred of the enemy but because they create a sense of expanded agency, invulnerability, and therefore, courage in battle.

<RB> R1.1. Self-sacrifice by fused individuals is motivated by egoism not altruism

The identity synergy theory of fusion (Swann et al. 2012), proposes that activation of personal identity in a fused individual also activates group identity and vice-versa. This would explain why fused individuals take any attack on the group personally, and therefore sacrifice self for group, as many past studies suggest. But it also raises the question whether it cuts both ways, such that fused individuals would show greater willingness to sacrifice group for self. Gaertner et al. summarize the results of recent studies, suggesting this is indeed the case, and supporting their view that the reason fusion increases willingness to fight and die for a group is not because of elevated altruistic feelings and diminished concern for self-preservation but because egoistic motivations are put into the service of group goals. This point is well taken, and future studies should seek to establish whether egoism, rather than altruism, motivates actual (as opposed to hypothetical), willingness of fused individuals to put their lives on the line to save the group, for example, in frontline armed combat or suicide missions. That said, even if the psychological motivation to sacrifice self for the group is egoistic rather than altruistic in fused individuals, the behavior would nevertheless still be describable as altruistic (i.e. benefitting others at cost to self).

<RA> R1.**On development**

<RB>R1.1. Fusion has its roots in early childhood

According to Rochat and also Thomsen & Fiske, the building blocks of identity fusion develop very early. Rochat traces fusion to conformism bias and sensitivity to ostracism in early childhood, as well as attentiveness to cues signaling socially salient group differences such as those based on gender, race, and economic status. Thomsen & Fisk trace the roots of altruism even further back, to infancy. There is indeed good evidence that, from an early age,

children show a strong desire to affiliate with groups and avoid exclusion. In a series of collaborations with developmental psychologists, we have found evidence that group identity markers, such as cultural rituals and conventions, motivate conformism and reduce independent innovation in pre-school children (Legare et al. 2015), while also leading them to cleave to local customs even more closely when primed with ostracism threats (Watson-Jones et al. 2014, 2016). But this is not the same as fusing with a group.

Nor can I agree with Rochat's claim that "dying for the group is nothing more than the extreme uncanny expression of the developing human need to affiliate and the deep fear of being rejected by those who provide basic support." Most children wish to join groups and avoid rejection by them but would not be willing to lay down their lives to protect the group against its enemies. Efforts to uncover the developmental pathways to fusion are only preliminary (Gaviria et al. 2015), but early evidence suggests that the ability to fuse with a group, in the way that adults do, actually emerges relatively late in development, during adolescence (Tasuji et al. in prep).

<RB>R1.1. Groups come before fusion

The idea that people can fuse with a group raises the question of what comes first, a group identity, or fusion with others, although it is also possible that these processes unfold together, interactively, or in some kind of feedback loop. Khalil argues that group identity must be established first, before a process of fusion can occur. It is not clear whether he means by this, that group identity is logically prior to fusion with a group, or if this is a developmental claim, in the sense that the capacity to identify with a group emerges earlier in life than the ability to fuse. Arguably, both claims have some prima facie plausibility. Much depends, however, on what one means by a "group." In the target article, a key distinction

was made between: (1) local groups based on relational ties (ranging from dyadic ties between siblings as in the case of our twins studies, through to larger networks of family-like bonds in a military unit or football team); (2) extended groups based on categorical ties (ranging from a territorially bounded country, to a globally distributed religious tradition, or ethnicity). As noted above, while children are sensitive to the presence of both local and extended groups from an early age, on current evidence, the ability to fuse with groups comes later. So, in that sense, Khalil may be right that groups come before fusion.

On the other hand, not all groups we fuse with are ones that we have grown up with. People commonly form or join groups in adolescence or adulthood, and while this could happen prior to fusing, that is not necessarily the case. For example, adult victims of a terrorist attack might have memories of shared suffering with others that subsequently lead to the formation of a highly fused group of survivors. Rather than insisting that groups always come before fusion, or that fusion always comes before groups, it seems more realistic to entertain both possibilities, as well as scenarios in which both develop in tandem.

<RB>R1.1. If fusion is a slow process, self-sacrifice should increase later in life (but doesn't)

As noted earlier, in response to commentaries suggesting that fusion is triggered instantaneously by negative events, the imagistic pathway to fusion model actually proposes a more gradual process of remembering, reflection, and identity transformation, that slowly and incrementally fuses personal and group selves. As Rosenweig et al. observe, however, if that account is correct, then surely fusion should increase in a linear fashion over the life cycle, with highest levels of fusion being apparent among the elderly who have had longer to

accumulate shared experiences and to reflect on them. Such an inference was not intended, however. In fact, initial efforts to test the "imagistic pathway" model longitudinally suggest that the process takes months or at most a few years (but not decades), to complete (Buhrmester et al. 2018; Muzzulini et al. in prep). This suggests that "trait fusion," a stable and potentially lifelong feeling of oneness with the group, may be assembled quite early life. Interestingly, this is the opposite of what dissonance theory would predict, namely that the longer people invest time and resources in a group, the more loyal its members will be. Both hypotheses were recently put head to head in a study of lifelong loyalty among football fans, supporting the hypothesis that fusion results from shared life-shaping experiences but that dissonance does not contribute to this causal chain (Newson et al. 2016).

In short, fusion would not be expected to increase steadily over a person's lifespan, and is more likely to emerge over a shorter timeframe and then stabilize. Still, Rosenweig et al. observe that extreme self-sacrifice is more likely to occur in adolescence or early adulthood and the framework proposed in my target article does not adequately explain why. They propose an intriguing answer to this question. Late adolescence is a period in which identity salience is especially high, potentially moderating the extent to which shared experiences produce fusion and to which already fused individuals are willing to die for the group. These proposals are consistent with existing evidence and well motivated theoretically, and, as such, should be the subject of future investigations.

<RA>R1.On function

<RB>R1.1. Social signaling explains extreme self-sacrifice

Efforts to demonstrate a biologically evolved function of extreme self-sacrifice often appeal to kin selection by arguing that even though heroes may take themselves out of the gene pool, relatives of heroes consistently accrue enough of a reproductive advantage for this to become an adaptive strategy over many generations. Dessalles argues that the demise of heroes benefits their kin by raising their social status. Heroes signal that they are good friends, especially in times of war. Being friends with a hero, presumably, makes people more likely to come under their protection. So, this type of "primary order social signaling" would explain why people want to befriend a hero. But how and why should their families benefit from this? Here, the argument becomes a little more complicated. If one cannot befriend a hero directly, then one should at least try to ingratiate oneself with their family and social network. This would explain why a hero's kith and kin are granted higher status. Maybe, but why then should this status persist after the hero dies? Again, Dessalles has an answer: This is due to "secondary order social signaling," whereby those honoring both heroes and their families derive status themselves by demonstrating that they are good patriots.

Each step of the social signalling argument raises thorny questions, however. Does it really make good evolutionary sense to seek the friendship of heroes? The implication is that friends of heroes benefit preferentially from their protection but one might equally argue that the point of heroism is that it isn't directed only to one's friends. Besides, hanging out with heroes might actually be quite dangerous, carrying an obligation to act bravely as well. Moreover, once the social signaling argument appeals to secondary order functions, it becomes even more tortuous. Given that the cost of honoring heroes may be low and the risk of insincerity high, in what sense is this a valid social signal at all? It is hard to see how this effort to explain the evolution of self-sacrifice improves upon the explanations already proposed in the target article. Whereas the latter have produced hypotheses capable of being

tested via both mathematical models and empirical research (Whitehouse & Lanman 2014; Whitehouse et al. 2017), the predictions of social signalling theory, in this context, are somewhat obscure.

<RB>R1.1. Extreme self-sacrifice is not an evolutionary puzzle

In contrast with Desalles, who tries to solve the puzzle of how extreme self-sacrifice could have evolved, Colman & Pulford argue that there is nothing puzzling to explain. According to the latter, self-sacrifice is only puzzling if one proceeds from the assumption that humans are basically selfish – or to use their term, "psychological egoists." But this assumption, they argue, is either tautological (since humans are selfish, every apparently altruistic act is really selfish ultimately), or false (contrary to empirical evidence that humans are naturally altruistic, or at least some are, under certain conditions). Colman & Pulford proceed to list various types of evidence that humans act altruistically and argue that we should not find this surprising. What remains unclear, however, is whether they actually intend to challenge the selfish gene hypothesis, according to which, self-sacrifice in the absence of selfish genetic benefits, cannot become an evolutionary stable strategy. None of the arguments or evidence presented actually constitutes a viable challenge to this hypothesis. Humans, like all other species, are indeed motivated to maximise access to resources that will quench their selfish appetites. But to the extent that cooperation can improve access to such resources, and solve the free-rider problem, cooperative behavior can evolve. This still does not explain extreme self-sacrifice, however. For such behavior to become biologically adaptive, it must result in the individuals who perform the behavior leaving more copies of their genes than those who do not. This is indeed a puzzle on the face of it, because dying is not an obviously good way

of passing on one's genes. That is why people like Desalles invoke kin selection as an explanation.

<RB>R1.1. Mutualism not kin selection explains extreme self-sacrifice

Strong forms of mutualism provide an elegant and compelling explanation for heroic behavior on the battlefield, based on the logic: I'll die for you, if you'll die for me, if (as a consequence) we both stand less chance of dying. Marie argues that costly signals of commitment in war could simply be scaled-up versions of the more general human propensity to secure assistance from others by signalling pro-sociality. The main difference is that on the battlefield, individual survival prospects depend more urgently on the protectiveness of one's fellows despite high personal risk, than in peaceful conditions. If the evolved function of heroism is to increase one's chances of survival then the death of a valiant fighter is always unintended. In principle, this account may be consistent with the conceptual framework presented in my target article, not only the imagistic pathway to fusion account (whereby the horrors of combat, for example, lead to heightened commitment to one's ingroup), but also the ultimate explanations proposed, since strong mutualism does not rule out other explanations (e.g. kin selection). Another version of this argument, which explicitly connects with kin psychology, is provided by Cronk's description of fitness interdependence, i.e. the extent to which one individual influences the survival and reproduction of another. Extreme self-sacrifice is most commonly observed in biological systems exhibiting high levels of fitness interdependence. Moreover, certain situations entail higher levels of fitness interdependence than others, for instance, cooperatively fending off a predator. In human societies, warfare would increase fitness interdependence, motivating

higher levels of self-sacrifice, perhaps recruiting kin psychology in the process. These are all interesting proposals that need to be tested.

<RB>R1.1. Fusion is not an adaptation for intergroup conflict

The idea that fusion evolved in conditions of widespread and chronic warfare, is questioned by Wiessner, based on the following considerations: First, the need to procure suitable mates would have necessitated intergroup marriage, and therefore, peaceful, rather than warlike, intergroup relations. I am not so sure. Methods of procuring mates in the ancestral past need not have been exclusively peaceful and the practice of bride capture is quite compatible with intergroup raiding and warfare. Second, lengthy childhood dependency encourages cooperative rearing practices, requiring in-group cohesion. Although apparently intended as an argument against the prevalence of intergroup conflict in ancestral societies, it is not clear why. Third, hunter-gatherers require large territories and a cooperative approach to facilitating access to each other. In support of this, Wiessner alludes to contemporary foraging societies famous for their peaceful egalitarianism but does not mention more warlike hunter-gatherers that until recently harbored imagistic warrior cults and high levels of intergroup conflict (Allen & Jones 2014; Pinker 2011). Fourth, warfare limits mobility "leaving belligerent groups few options." The logic here seems to be that warlike foraging bands would have had more limited opportunities to access the resources of their more peaceful neighbors. But surely the opposite is likely to have been true: More assertive groups willing to resort to violence, would have had better access to resources than those inclined to appease or flee.

<RA>R1.On history

<RB>R1.1. Historical traumas and not just personally experienced ones can motivate fusion

Babińska makes the important point that if shared suffering can be experienced vicariously, and not just directly, this means that historical traumas and other group-defining events in the past could drive fusion in subsequent generations. Babińska argues that this process operates at the level of extended, rather than local groups but that local fusion based on directly-shared experiences in the present can in turn strengthen reciprocally the connection to ancestral struggles. While we might disagree about the details of how this could work (e.g. Babińska equates extended fusion with strong identification, whereas I regard these as quite distinct constructs), it is surely an interesting question about the way in which fusion is generated and passed on within groups.

Arguably, the distinction between vicariously and directly shared experiences need not correspond in any simple way to the distinction between extended and local fusion. Indeed, one could share experiences vicariously via relational ties in a local group, such as family (e.g. children adopting the group alignments of their parents based on empathetically reliving the latter's experiences), and not just in an extended group (e.g. the descendants of an oppressed religious sect reexperiencing the persecutions of their forebears). At any rate, the broader point is that history may heavily influence group alignments in the present and give rise not only to fusion – whether extended or local – but also to perceptions of out-group threat, potentially fueling extreme self-sacrifice among generations to come. If that is the case, this is an issue of high importance not only scientifically but potentially also from a policy perspective.

<RB>R1.1. Relational ties precede groups in the prehistory of self-sacrifice

Palmer and Clark argue that groups cannot be regarded as relevant entities in explaining self-sacrificial behavior:

> "It is neither necessary nor accurate to describe self-sacrifice as being performed for some "group," nor caused by fusion with that "group." The ethnographic record describes the gatherings of foragers as fluid, with individuals who are kin gathering, dispersing, and regathering in different combinations. Although the fluidity of "groups" is often claimed to be compatible with multi-level selection that includes selection at the level of "groups," surely there is some degree of fluidity past which claims of individuals being divided into "groups" becomes false."

This passage seems to suggest that Palmer and Clark do not believe humans align with and act as groups, or at least that foraging societies do not. Setting aside the risks of generalizing from contemporary acephalous hunter-gatherers to prehistoric societies, the argument that people do not fuse with (or presumably identify with or even recognize the existence of) groups is somewhat baffling, given the weight of empirical evidence to the contrary. If what Palmer and Clark mean is simply that the groups people believe they are dying for when they go into battle are actually mental constructs rather than objective features of the world, that may be true in some (obscure) sense but is it relevant? Arguably, what matters for our purposes, is that people do indeed align with entities they construe to be groups, as well as with relational ties, and sometimes these alignments are so strong they are willing to fight and die to protect those in-groups against out-groups. Reconstructing the history of these behaviors is no easy task, requiring triangulation across multiple sources of theory and

evidence – not only ethnography (to which Palmer and Clark appeal), but also such disciplines as archaeology, history, evolutionary anthropology, and primatology. One of the most serious challenges facing this endeavor, however, is the temptation to cherry pick cultural traits or features of social morphology and assemble them into just-so stories about the evolution of human civilizations. To avoid this problem, we need to use evidence about the past in new ways, as argued in the ensuing subsection.

<RB>R1.1. To explain the cultural evolution of extreme self-sacrifice historically we must overcome selection bias

Although Tinbergen's fourth question, pertaining to the shaping and constraining effects of history in the evolution of traits, is just as important as the other three, it attracted the least sustained attention in the commentaries. Nevertheless, my target article advanced a series of specific claims about the way local fusion and its capacity to motivate extreme self-sacrifice has been shaped historically, with the evolution of social complexity. For example, I argued that while highly fused "bands of brothers" have always been a recurrent feature of small-scale societies engaged in high-risk pursuits such as raiding, warfare, and large game hunting, the rise of states and empires tended to outlaw and marginalize such groups, as a threat to centralized authority, or else harnessed the power of local fusion in highly restricted and carefully controlled environments such as the military and elite institutions. This is a testable claim but also one that has yet to be investigated systematically. Efforts to explore the evidence from history and archaeology focusing on case studies (Whitehouse and Martin 2004; Martin and Whitehouse 2005), have always been vulnerable to the charge of cherry picking examples that fit the theory. The challenge is to devise a method of testing theories of

the evolution of self-sacrifice more objectively, especially in ways that overcome the problem of selection bias. I therefore direct this criticism to myself, and attempt to address it.

One solution to the problem of selection bias is to quantify patterns in human history, while controlling for nonindependence, so that we can test competing hypotheses systematically. The target article introduced one such ambitious effort known as Seshat: Global History Databank (Turchin et al. 2015; Turchin et al. 2018; Whitehouse, 2016). A promising feature of this approach is that it will enable researchers to test, not only the above predictions about the fate of imagistic groups, but also a range of alternative theories, including those that make even more precise or nuanced predictions about the role of group psychology in extreme self-sacrifice. For example, the "local fusion plus threat" model of self-sacrifice proposed here, might be better able to explain the ebb and flow of specific practices such as suicide terrorism, by adding just one or a few additional variables into the mix (e.g. histories of appeasement on the part of ruling groups).

<RA>R1.Extending the framework

As well as critiquing the conceptual framework advanced in the target article, some of the commentaries propose ways of extending it in novel directions. The boundary between critiquing and extending can be blurry, of course. For example, some of the criticisms fielded above, also suggest the possibility of extending the framework by taking additional variables into account. Nevertheless, in this section of the response, I consider commentaries that are primarily seeking to add to, rather than to contest, the conceptual framework, as presented. In terms of Tinbergen's four questions, all commentaries treated in this final part of the response are primarily concerned with problems of proximate causation.

[COMP: INSERT TABLE 2 with Table 2 Caption HERE]

<RB>R1.1. Entitativity explains why terrorists target the innocent and inherit grievances

Even if the fusion plus threat formula explains willingness to lay down one's life to defend the group, it would not explain why so many terrorist attacks target the innocent. Choi et al. argue that this is due to the attribution of entitativity to out-groups, such that all their members are seen as equally culpable for past offenses and therefore, equally suitable as targets for revenge. Moreover, they suggest that this same logic of entitativity would explain why grievances often get passed down through the generations, and why armed groups may claim to be defending an ancestral group as if its members were still alive today. Considering first the question why terrorists target the innocent, a plausible alternative explanation is that they do so for tactical reasons, for example, in order to frighten the population at large and thereby pressurize governments to accede to their demands (Pape 2006). If killing and maiming the innocent is primarily a tactic, then entitativity may contribute mainly to the construction of post hoc rationales rather than a desire to harm bystanders. This could explain why some members of terrorist organizations, such as former Jemaah Islamiya member Nasir Abbas, have disputed the rightness of targeting civilians while still defending the right to kill out-groups bearing arms.

The inheritance of grievances over the generations also raises interesting questions. While entitativity may well play a role, another factor worth considering may be vicarious episodic memory. It is possible that some people identify so closely with the narratives of

parents, grandparents, or even more ancient ancestors that their sufferings come to be functionally equivalent to personal experiences in autobiographical memory. If so, vicarious shared experiences could form a basis for fusion with a historical group, just as with a contemporary one. This possibility, as well as the entitativity proposal, suggests exciting new directions for research on extreme self-sacrifice.

<RB>R1.1. Violence-condoning norms contribute to violent extremism

Louis et al., Elnakouri et al. and Ginges and Shackelford point out that not all fused groups that come under attack respond with violence, and those that do typically subscribe to violence-condoning norms (Newson et al. 2018). Elnakouri et al. add a further interesting layer to this line of argument by suggesting that violence-condoning norms arise, and are more enthusiastically endorsed, in response to environmental threat. In addition to looking more closely at the role of norms, Louis et al. identify other variables that may contribute to fused groups turning to violence in the face of out-group threats, such as the absence of peaceful means of defending the group's interests. To this, we might add, that access to weaponry, such as guns and explosives, could increase the risk of fused groups turning to violence in the face of threat, although we also have many examples of suicide terrorists using everyday objects (e.g. ranging from knives to vehicles), rather than military-grade weapons, to carry out attacks. As the conceptual framework evolves, it will probably be necessary to add new mediations and moderated mediations but the costs in complexity will need to be weighed against empirical gains, as determined by future research.

<RB>R1.1 Individual differences predict propensity to fuse and sacrifice self

Stagnaro et al. and Rahal point out that not only situational factors but also individual differences could help to explain extreme self-sacrifice. Stagnaro et al. argue that individuals who have a more intuitive rather than deliberative cognitive style, may be more likely to fuse with a group and to act impulsively on the urge to protect it against out-group attacks, an observation that finds some support in empirical research (Fredman et al. 2015). Further, they suggest that other individual differences, for example, with regard to attachment, aggression, and empathy, could also increase willingness to sacrifice self for group. These suggestions seem worth investigating further but, as Stagnaro et al. also recognize, the process of fusing is often highly deliberative rather than intuitive (Jong et al. 2015). In a somewhat similar vein, Rahal argues that individual level tendencies toward self-harming behavior as well as willingness to join in intergroup conflicts, might both increase the probability that somebody would be willing to fight and die for a group. These suggestions are all well taken. Nevertheless, it should be noted that early efforts to investigate whether identity fusion taps individual differences – including empathy, aggressiveness, self-efficacy, self-concept clarification, and essentialism – found no evidence for this (Gomez et al. 2011). Moreover, while one can see why differences in aggressiveness might be thought to affect levels of out-group hostility, it is not clear how they relate to willingness to sacrifice self for group, one of the key issues under consideration here.

<RA>R1.Conclusions and future directions

Taking the commentaries as a whole, it is possible to distinguish six clusters of questions that warrant further investigation. One cluster concerns the role of belief. For example, do beliefs motivate extreme self-sacrifice independently of identity fusion? Or alternatively, does sharing a set of core values constitute an additional pathway to fusion (alongside shared

experience and shared biology)? Is threatening a group's beliefs equivalent psychologically to threatening its individual members?

A second cluster concerns the psychology of fusion itself. When highly fused individuals act to defend their groups, are they motivated by egoism (e.g. self-preservation) or altruism (e.g. concern for the welfare of others)? What can studies of social synchrony tell us about the mechanisms underlying identity fusion? How long does it take for a shared experience to increase trait fusion and what exactly is the process?

A third cluster concerns changes in fusion over the life cycle. For example, do different pathways to fusion emerge during different periods of development (e.g. does the shared biology pathway become established earlier than the shared experience one)? Does group psychology emerge prior to the capacity to fuse or the other way around? Or neither? Why are fused adolescents, and especially male adolescents, more likely to fight and die for a group?

A fourth cluster focuses on the nature and diversity of fusion targets. For example, how does fusion with a relational group differ from fusion with a group category? How far can fusion be extended? Can one fuse with anyone or anything? To what extent does moral expansiveness affect the degree to which fusion can be extended to encompass species and even objects and concepts unrelated to any in-group?

A fifth cluster of questions concerns the nature of the link between fusion and self-sacrifice. In additional to out-group threat, several potential moderators of the relationship between fusion and violent self-sacrifice were suggested, including: violence-condoning

norms, access to deadly weapons, the absence of peaceful mechanisms for conflict resolution, entitativity, segregation of in-groups and out-groups, and fitness interdependence among in-group members. To what extent do individual differences (e.g. with regard to impulsivity and aggression) affect the likelihood of fusion to motivate violent self-sacrifice? And does fusion predict fighting or dying equally, or one more than the other?

Finally, a sixth cluster concerns the extent to which group experiences can be transmitted across generations. Can the effects of shared experiences and out-group threat on warfare intensity be detected in the human past? If so, can historical patterns of group bonding and warfare be used to predict the outcomes of present and future conflicts?

I would like to thank all the commentators for generating this formidable list of questions and to apologize if any have been inadvertently neglected. Taking these questions seriously and building new research designs to investigate them, will not only enable us to understand better why people can be so passionately committed to their groups but may also help us develop interventions to make the human world more peaceful.

<R-Text ends>

<RFT>References [Harvey Whitehouse] [rHW]

<REFS>

M.W,Allen & T.L. Jones ed. (2014). *Violence and warfare among hunter-gatherers*. Left Coast Press [rHW]

Atran, S., Sheikh, H. & Gómez, Á. (2014). Devoted actors fight for close comrades and

sacred cause. *Proceedings of the National Academy of Sciences, USA.* 111:17702-17703. [rHW]

Buhrmester, M. D. (2013). Understanding the cognitive and affective underpinnings of whistleblowing. *University of Texas ScholarWorks. Available at:* http://hdl.handle.net/2152/21278 [rHW]

Buhrmester, M.D. & Lanman, J. (in prep). For cause or kin? Identity fusion predicts self-sacrifice for one's religion. [rHW]

Buhrmester, M., Burnham, D., Johnson, D., Curry, O.S., Macdonald, D. & Whitehouse, H. (2018). How moments become movements: Shared outrage, group cohesion, and the lion that went viral. *Frontiers in Ecology and Evolution* Vol.6. doi:10.3389/fevo.2018.00054 [rHW]

Buhrmester, M. D., Brannon, S.M., Fraser, W.T., Fredman, L.A., Gomez, A., Talaifar, S. & Swann, W. B., Jr. (2015). Identity fusion, extreme pro-group behavior, and the path to defusion. *Social and Personality Psychology Compass* 9(9): 468–80, doi:10.1111/spc3.12193 [rHW]

Ferreira, C., Gaviria, E., Martínez, M. and Whitehouse, H. (2015). Identity and the developmental origins of fusion: An exploratory approach / La identidad y los orígenes de la fusion en el desarrollo: un enfoque exploratorio. *Revista de Psicología Social / International Journal of Social Psychology* 30(3): 531-62. doi:10.1080/02134748.2015.1065088 [rHW]

Gomez, A.,Brooks, M.L., Buhrmester, M.B., Jetten, J., Swann Jr. W.B. & Vazquez, A. (2011). On the nature of identity fusion: Insights into the construct and a new measure. *Journal of Personality and Social Psychology.* 100 (5): 918–933. Doi:10.1037/a0022642 [rHW]

Gómez, A., Atran, S., Davis, R., Ginges, J., López-Rodríguez, L., Sheikh, H., Waziri, H., Wilson, L. & Vázquez, A. (2017). The devoted actor's will to fight and the spiritual dimension of human conflict. *Nature Human Behaviour* 1:673-79. [rHW]

Halberstadt, J., Bilkey, D., Jackson, J.C., Jong, J., McNaughton, C., Whitehouse**,** H. & Zollmann, S. (2018). Synchrony and physiological arousal increase cohesion and cooperation in large naturalistic groups. *Nature: Scientific Reports.* 127. doi:10.1038/s41598-017-18023-4 [rHW]

Jong, J., Kavanagh, C., Lane, J. & Whitehouse**,** H. (2015). Shared negative experiences lead to identity fusion via personal reflection. *PloS ONE* doi: 10.1371/journal.pone.0145611 [rHW]

Jong, J., Kavanagh, C. M., McKay, R. & Whitehouse, H. (2018). Positive experiences of high arousal martial arts rituals are linked to identity fusion and costly progroup actions. *European Journal of Social Psychology.* doi:10.1002/ejsp.2514 [rHW]

Legare, C. H., Herrmann, P.A., Wen**,** N.J. & Whitehouse, H. (2015). Imitative flexibility and the development of cultural learning. *Cognition*. 142:351-61. doi: 10.1016/j.cognition.2015.05.020 [rHW]

Muzzulini et al., Kapitany, R., van Mulukum, V. & Whitehouse, H. In Prep [rHW]

Peregrine, P. N. Brennan, R., Currie, T., Feeney, K., François, P., Turchin, P. & Whitehouse, H. (2018). Dacura: A new solution to data harvesting and knowledge extraction for the historical sciences. *Historical Methods: A Journal of Quantitative and Interdisciplinary History. doi:10.1080/01615440.2018.1443863* [rHW]

Pinker, S. (2011). *The better angels of our nature*. Viking. [rHW]

Reddish, P., Jong, J., Lanman, J. A., Tong, E. M. W. & Whitehouse, H. (2016). Collective synchrony increases prosociality towards non-performers and outgroup members. *British Journal of Social Psychology*. doi:10.1111/bjso.12165 [rHW]

Reddish, P., Jong, J., Tong, E. & Whitehouse, H. In Prep. Interpersonal synchrony affects performers' sense of agency. [rHW]

Rybanska, V., Jong, J., McKay, R. & Whitehouse, H. (2017). Rituals improve children's ability to delay gratification. *Child Development.*
89 (2): 349-59. doi.org/10.1111/cdev.12762. [rHW]

Swann, W. B., Bastian, B., Gómez, A., Jensen, J. & Whitehouse, H. (2012). When group membership gets personal: A theory of identity fusion. *Psychological Review* 119(3): 441–56. [rHW]

Swann Jr, W. B., Bastian, B., Buhrmester, M. D., Finchilescu, G., Gómez, A., Jetten, J. & Vázquez, A. (2014). What makes a group worth dying for? Identity fusion fosters perception of familial ties, promoting self-sacrifice. *Journal of Personality and Social Psychology* 106(6):912. [rHW]

Tasuji, T., Whitehouse, H., Chan, J., Henderson, D., & Reese, E. In Prep. The development of identity fusion in adolescence via shared experiences. [rHW]

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, 20: 410–33. [rHW]

Turchin, P., Brennan, B., Currie, T.E., Feeney, K.C., François, P., Hoyer, D., Manning, J.G., Marciniak, A., Mullins, D., Palmisano, A., Peregrine, P., Turner, E.A.L. & Whitehouse, H. (2015). Seshat: The global history databank. *Cliodynamics: The Journal of Quantitative History and Cultural Evolution* 6(1): 77-107. [rHW]

Turchin, P., Currie, T.E. & Whitehouse, H. et al. (2018). Quantitative historical analysis uncovers a single dimension of complexity that structures global variation in human social organization. *Proceedings of the National Academy of Sciences of the United States of America (PNAS).* 115(2): E144-E151. [rHW]

van Mulukom, V., Muzzulini, B., Whitehouse, H. &Willard, A. Autobiographical memory pathway from shared dysphoria to identity fusion. In prep. [rHW]

Vázquez, A., Gómez, A., Ordoñana, J.R., Swann, W.B. & Whitehouse, H. (2017). Sharing genes fosters identity fusion and altruism, *Self and Identity* 16:6:684-702. doi:10.1080/15298868.2017.1296887.

Watson-Jones, R., Clegg, J., Legare, C.H. & Whitehouse, H. (2014). Task-specific effects of ostracism on imitation of social convention in early childhood. *Evolution and Human Behavior*. 35(3): 204 – 10. [rHW]

Watson-Jones, R.E., Legare, C.H. & Whitehouse**,** H.. (2016). In-group ostracism increases high fidelity imitation in early childhood. *Psychological Science* 27(1): 34–42. doi: 10.1177/0956797615607205 [rHW]

Whitehouse, H. (1992). Memorable religions: Transmission, codification, and change in divergent Melanesian contexts. *Man* (N.S.) 27(3): 777-97. [rHW]

Whitehouse, H. (2004). *Modes of religiosity: A cognitive theory of religious transmission*. AltaMira Press. [rHW]

Whitehouse, H. (2016). Ritual and social evolution: Understanding social complexity through data. In B. Bozic et al. eds. *Computational History and Data-Driven Humanities*. doi: 10.1007/978-3-319-46224-0 1 [rHW]

Whitehouse, H. & Lanman, J.A. (2014). The ties that bind us: Ritual, fusion, and identification. *Current Anthropology* 55(6): 674 – 95. [rHW]

Whitehouse, H., Buhrmester, M., McQuinn, B. & Swann, W. B. (2014). Brothers in aArms: Libyan revolutionaries bond like family. *Proceedings of the National Academy of Sciences*. 111(50): 17783-17785. Early Edition www.pnas.org/cgi/doi/10.1073/pnas.1416284111. [rHW]

Whitehouse, H., Bastian, B., Buhrmester, M.D., Gavrilets, S., Gómez, A., Jong, J., Kavanagh, C.M., Lanman, J.A., Matthews, M., McKay, R. & Newson, M. (2017). The evolution of extreme cooperation via intense shared experiences. *Nature: Scientific Reports.* doi:10.1038/srep44292. [rHW]

<TABLE 1>

Tinbergen's question on:	**Criticism**	**Commentator**
Mechanism	Shared values and beliefs can motivate extreme self-sacrifice	Atran & Gómez; Swann & Jetten; Boudry; Ginges & Shackleford; Kiper & Sosis; Lane et al
	Communal sharing rather than fusion explains self-sacrifice	Thomsen & Fiske
	Segregation not fusion increases the effects of out-group threat on extreme self-sacrifice	Kiper and Sosis; Melton and Motyl
	Self-sacrifice is not always about in-groups	Crimston & Hornsey; Elnakouri et al
	Fusion does not always lead to violent extremism	Kiper & Sosis;

		Xygalatas; Hansen
	Shared trauma can trigger group bonding instantaneously	Lankford; Swann & Jetten; Xygalatas
	Self-sacrifice by fused individuals is motivated by egoism not altruism	Gaertner et al
Development	Fusion has its roots in early childhood	Rochat; Thomsen & Fiske
	Groups come before fusion	Khalil
	If fusion is a slow process, self-sacrifice should increase later in life (but doesn't)	Rosenweig et al
Function	Social signalling explains extreme self-sacrifice	Dessalles
	Extreme self-sacrifice is not an evolutionary puzzle	Colman & Pulford
	Mutualism not kin selection explains extreme self-sacrifice	Marie; Cronk
	Fusion is not an adaptation for intergroup conflict	Wiessner
History	Historical traumas and not just personally experienced ones can motivate fusion	Babińska
	Relational ties precede groups in the prehistory of self-sacrifice	Palmer & Clark
	To explain the cultural evolution of extreme self-sacrifice historically we must overcome selection bias	Whitehouse

<TABLE 1 CAPT>

Table 1. Tinbergen's questions applied to extreme self-sacrifice

<TABLE 2>

Extension proposal	**Commentator**
Entitativity explains why terrorists target the innocent and inherit grievances	Choi, Jackson & Gelfand
Violence-condoning norms contribute to violent extremism	Ginges & Shackleford; Elnakouri et al. Louis et al.
Individual differences predict propensity to fuse and sacrifice self	Stagnaro, Littman & Rand; Rahal

<TABLE 2 CAPT>

Table 2. Proposed extensions to the conceptual framework