



# The Scapegoat Mechanism in Human Evolution: An Analysis of René Girard's Hypothesis on the Process of Hominization

D. Vincent Riordan<sup>1</sup>

Received: 15 September 2020 / Accepted: 19 April 2021 / Published online: 22 July 2021  
© The Author(s) 2021

## Abstract

According to anthropological philosopher René Girard (1923–2015), an important human adaptation is our propensity to victimize or scapegoat. He argued that other traits upon which human sociality depends would have destabilized primate dominance-based social hierarchies, making conspecific conflict a limiting factor in hominin evolution. He surmised that a novel mechanism for inhibiting intragroup conflict must have emerged contemporaneously with our social traits, and speculated that this was the tendency to spontaneously unite around the victimization of single individuals. He described an unconscious tendency to both ascribe blame and to imbue the accused with a sacred mystique. This emotionally cathartic scapegoat mechanism, he claimed, enhanced social cohesion, and was the origin of religion, mythology, sacrifice, ritual, cultural institutions, and social norms. It would have functioned by modifying the beliefs and behaviors of the group, rather than of the accused, making the act of accusation more important than the substance. This article aims to examine the empirical evidence for Girard's claims, and argues that the scapegoat hypothesis has commonalities with several other evolutionary hypotheses, including Wrangham's execution hypothesis on self-domestication, Dunbar's hypothesis on the role of storytelling in maintaining group stability, and DeScioli and Kurzban's hypothesis on the role of non-consequentialist morality in curtailing conflict. Potential implications of the scapegoat hypothesis for evolutionary psychology and psychiatry are discussed.

**Keywords** Aggression · Girard · Mentalizing · Religion · Sacrifice · Scapegoating

## Introduction

Humans live in a complex social environment characterized by large group sizes and cooperation between relative strangers. Uniquely human traits that have been suggested as being of importance to this include our abilities to mentalize (have a theory of mind) (Premack and Woodruff 1978), a related tendency to create a shared attentional framework or a shared intentionality (Tomasello and Carpenter 2007), a relative inhibition of reactive aggression (Wrangham 2019a), and the emergence of behavioral norms with an associated propensity to punish transgressors (Tomasello 2016). Some of these traits have, variously, been linked to the emergence of relatively large human group sizes (Dunbar 1998),

complex civilizations (Norenzayan et al. 2016), and uniquely human hierarchical structures (Dubreuil 2010).

This article considers the claim that another important and uniquely human psychological trait is that of victimization or scapegoating. This was the view proposed by the French anthropological philosopher René Girard (1923–2015) who posited that a propensity to internecine violence was an inevitable consequence of some of the otherwise fitness-enhancing human social traits (Girard et al. 1978; Girard 1986). He proposed that a novel mechanism must have emerged which countered this destabilizing influence, and that conspecific conflict in early hominins was curtailed by the emergence of a propensity to respond to adversity by blaming arbitrarily selected individuals. He called this the “scapegoat mechanism,” claiming that such victimization, coupled with a tendency to retrospectively misunderstand this pacification process and the role played by the victim, became the foundation of myths, religion, ritual, culture, and social norms.

✉ D. Vincent Riordan  
vincent.riordan@hse.ie

<sup>1</sup> West Cork Mental Health Services, HSE South, Bantry Hospital, County Cork, Ireland

This article aims to examine the evidence for Girard's claim that scapegoating was an adaptation of critical importance to the process of hominization by comparing it with other approaches to human evolution, and explores some potential implications for evolutionary psychology and evolutionary psychiatry.

## René Girard and the Scapegoat Mechanism

### Background

Girard advanced a wide-ranging theory of human nature and human origins (Girard 1961, 1972; Girard et al. 1978; Fleming 2004; Antonello and Gifford 2015). He has been described by his admirers as “the Darwin of the social sciences” (Serres 2009), although he had no formal background in science, but was a historian by training, who spent much of his academic career working in university language departments (Haven 2018). His hypothesis, described below, was born, not from empirical data as such, but rather from his interpretation of great works of literature, such as those of Shakespeare, Proust, and Dostoevsky (Girard 1961), as well as mythology (Girard 1972) and religious texts (Girard et al. 1978). He did take some inspiration from the field of anthropology (Frazer 1913; Evans-Pritchard 1962), although some of the conclusions he drew from anthropological data have been questioned (Traube 1979).

His work was thus born of the humanities rather than the sciences, and this may account for why he is not very well known amongst evolutionary theorists. Yet, although what he postulated is speculative, it is nevertheless a hypothesis concerning human evolution (Girard et al. 1978, pp. 84–104), and despite his own lack of formal scientific training, he was a polymath who considered himself to have been working within a Darwinian framework (Girard 2017).

His work is often described as “mimetic theory,” reflecting the central role he claimed for an innate human propensity to mimic or to imitate. The scapegoat mechanism, which is the subject of this article, comprises a major part of this theory, whereby he argued that humans evolved a tendency, when faced with a crisis, to spontaneously imitate each other in blaming, or scapegoating, arbitrarily selected individuals or groups (Girard 1986).

### The Posited Problem; Mimetic Rivalry and Violence

Before describing the important role claimed for scapegoating, it is necessary to first consider the selection pressures that Girard envisaged early hominins to have encountered, and to which he claimed scapegoating emerged as an adaptation. In other words, it is necessary to consider the uniquely

human problem to which scapegoating might have emerged as an adaptive solution.

As indicated above, the central plank of Girard's ideas is the claim that one of the most characteristic, if not defining, human traits is a tendency to mimic both the actions and the intentionality (mental states) of each other, a trait that he called “mimesis.” His primary focus was on what he called “mimetic desire,” and how the reciprocal imitation of desires can result in a positive feedback escalation, such that human desire could become abstract and detached from the functional utility of the object of desire. He considered how reciprocal mimesis can give desire, and other intentional states, a contagious quality, causing them to spread rapidly throughout a population. In collaboration with psychiatrists Jean Michel Oughourlian and Guy Lefort, he used this concept of reciprocal interdependent desires to posit what they called an “interdividual psychology” (Girard et al. 1978, pp. 283–431), which they used to propose explanations for various human behaviors and psychopathologies (Oughourlian 2016).

The focus of this article, however, is not so much on the claimed mimetic or imitative nature of human behavior, but rather on Girard's ideas regarding the resultant problems caused by human rivalry, feuding, and violence, and, in particular, on the solution, scapegoating, which he theorized as having emerged in response to this.

Although recognizing that the spontaneous alignment of the desires, motivations, and intentions of an entire group would have facilitated cooperation and efficiency, Girard also saw a problem. If everyone desires the same thing, then rivalries are inevitable. He considered this problem in terms of the dominance-based social hierarchies that seem to curtail conspecific violence in other animals (Girard et al. 1978, p. 95). The stability of such hierarchies depends on subordinates maintaining a different intentional state to dominants, hence any tendency for the mental states of all members of a community to align would inevitably destabilize such hierarchies, thereby undermining their ability to suppress conflict.

Therefore, he surmised, if a population evolves contagious mimetic behavioral traits, as humans would seem to have done, once conflict begins, it can quickly spread throughout a community. If good relations become reciprocal, then bad relations become reciprocal also. In other words, Girard suggested that the behavioral trait (mimesis) that was essential to humans becoming such an adaptable, cooperative, and successful species, is also the trait responsible for a more problematic human phenomenon, that of cyclical revenge. He theorized that the risk of “runaway” violence amongst early hominin groups would have threatened their survival, implying that it would have been a limiting factor to human evolution. He called such “all against

all” scenarios “mimetic crises,” which he said were characterized by a lack of difference between group members.

He claimed that mimesis would have undermined the intragroup differences upon which dominance-based hierarchical stability depended, or as he put it: “Beyond a certain threshold of mimetic power, animal societies become impossible” (Girard et al. 1978, p. 95). He surmised therefore, that as mimesis emerged there would have been a strong selection pressure for the contemporaneous emergence of some other behavioral trait that would have countered its inherently destabilizing influence on social hierarchies.

### The Posited Solution; the Scapegoat Mechanism

Between what can be strictly termed animal nature on the one hand, and developing humanity on the other, there is a true rupture, which is collective murder, and it alone is capable of providing for kinds of organisation, no matter how embryonic, based on prohibition and ritual.

(Girard et al. 1978, p. 97)

As the name implies, the scapegoat mechanism would have involved all members of a community blaming a single individual for whatever crisis or adversity was at hand. This would have had two effects. The most immediate consequence would have been to instill communal unity where previously there would have been conflict. The Hobbesian “all against all” (Hobbes and Oakeshott 1946) would have become the Girardian “all against one.” Mimetic crises, he envisaged, would have been resolved by all members of a group spontaneously aligning their hostility towards a single individual, and either expelling that individual from the community or, more probably, killing them. Such crowd behaviors would themselves have been products of mimesis. They are recognizable in modern humans in behaviors ranging from school yard bullying to mob lynching.

Of equal, if not greater, importance, Girard theorized, would have been a longer-term effect whereby this cathartic violent resolution to a crisis would have engendered a narrative, or myth, that initially blamed the selected victim for the crisis, but subsequently may have credited them with being the source of its resolution, imbuing them with a mystique, and even making them appear godlike. The community would have failed to recognize that it was their own actions and unity, rather than any attribute or action of the victim, that had resolved the mimetic crisis. Girard referred to this invariable misunderstanding as “misrecognition.” Paraphrasing psychoanalysis, he coined the term “double transference” to describe the tendency to project onto the victim both the blame for initiating the crisis, and the credit for resolving it.

This retrospective deification of the victim, he claimed, was the origin of the concept of “the sacred,” of archaic religion, and of mythology. He argued that hominin communities, when faced with future crises, would have tried to repeat, or reenact, the process by which an earlier mimetic crisis had been resolved. However, as they would not have understood the interpersonal dynamics that had caused (and resolved) the earlier mimetic crisis of internecine violence, they would have been unable to distinguish such an event from other crises, such as ones caused by natural disasters. Therefore, there would have emerged a tendency to reenact victimization as the default response to all crises, regardless of their actual cause or nature. This urge to reenact the originally cathartic scapegoat resolution, he claimed, was the origin of ritual, including human sacrifice, and funerary practices. Such rituals and shared beliefs would in turn not only have further curtailed intragroup violence, but would have positively enhanced communal unity.

Girard claimed that the various prohibitions associated with religious practices can mostly be understood as having originated as attempts to thwart rivalry and thus prevent further mimetic crises. These would have relied for their efficacy on shared communal beliefs, norms, and values, which in turn would have been intrinsically linked to the intergenerational transmission of unanimous belief in the misrecognition generated by the scapegoat mechanism.

According to Girard, myths are misunderstood accounts of real victimization events narrated through the prism of misrecognition and double transference. This misunderstanding would have been essential for the scapegoat mechanism to work effectively and generate unanimously accepted values and beliefs. However, he argued that the defining characteristic of modern societies is a realization of the probable innocence of the victim, a realization that he attributed to the pervasive influence of Judeo-Christianity and which, he claimed, has a desacralizing effect, making it more difficult for modern societies to generate, or even to maintain, unanimously accepted social norms.

Whereas the modern usage of the term scapegoating might imply a deliberate act on the part of a guilty party to divert blame elsewhere, Girard used it somewhat differently. Although common to both uses of the word is the implication that the accused is essentially innocent, Girard envisaged a more unconscious process, essentially a cognitive bias, not only to ascribe blame, but also to imbue the one accused with mystique.

He speculated that kingship had its origins in victimization, specifically in delayed victimization (Girard et al. 1978, pp. 55–57). As victims would have been perceived as having an ambiguous status, being simultaneously both malign and benign, if a sacrificial victim were to have been selected but not immediately killed, then during the intervening period

they would have been perceived as having had a sacred aura. This ambiguity, between king or shaman and sacrifice victim, may appear somewhat counterintuitive, but it is consistent with the idea that the scapegoat mechanism functioned by reestablishing (social) differences in a community where hierarchical differences had been erased by mimetic behaviors. Girard argued that the spontaneously generated binary difference between the victim and the crowd allowed a novel and uniquely human social order to emerge. Indeed, he claimed that ultimately all cultural forms have their roots in victimization and scapegoating. The mystique retrospectively attached to victims caused them to become what he called, “transcendental signifiers,” and thus the basis of symbolism, of meaning, and eventually of complex cultural institutions, including kingship and state power. This would have allowed for “sacred violence,” that is, violence sanctioned by a sacred king, which could be carried out with minimal risk of it precipitating reprisal or further violence.

To conclude, Girard advanced an ambitious, if somewhat speculative, hypothesis on human origins, but his critics accuse him of selective interpretation of literature and anthropology (Landy 2012; Merrill 2017). The following sections aim to appraise his scapegoat hypothesis in the light of other relevant hypotheses and empirical data, most of which have emerged in the decades after Girard formulated his ideas.

### Comparative Analysis; the Posited Problem; Mimetic Rivalry and Violence

Before discussing what evidence might or might not exist for the proposed solution, scapegoating, we first need to consider the plausibility of the problem. Would mimesis, or analogous social traits, have precipitated internecine conflict, thereby making “animal societies impossible”?

Girard considered the stability of “animal societies” in terms of dominance hierarchies, on the premise that such hierarchies have an important pacifying function in other animals. Although there would seem to be little direct empirical evidence to support this premise *per se*, other authors have made similar assumptions (Price 1967) and the interpersonal dynamics that result in a mimetic crisis are analogous to Gregory Bateson's concept of symmetrical differentiation causing schismogenesis leading to “hostility and the breakdown of the entire system” (Bateson 1935).

It seems plausible that hierarchical stability might have been disrupted by the emergence of at least two of the traits upon which human sociality depends. These are the ability to mentalize and the associated propensity to create a joint attentional framework or a “shared intentionality.”

### Mentalizing

As with much of human sociality, a prerequisite for Girard's concept of mimetic desire would have been the ability to mentalize, that is to accurately assess the mental states, or intentionality, of others. This ability has been described in terms of a hierarchy of orders of intentionality (Dennett 1989; Dunbar 2008), whereby first-order intentionality entails an awareness of one's own mental state; second-order intentionality, which is also referred to as “theory of mind” (Premack and Woodruff 1978; Baron-Cohen 1995), requires an ability to conceptualize the mental state of another (e.g., I believe that you believe); and, in turn, third-order intentionality refers to the ability to conceive of another's belief about another's mental state (e.g., I believe that you believe that I believe). There is, at most, weak evidence that some apes, such as chimpanzees, may have a limited ability for second-order intentionality (O'Connell and Dunbar 2003; Call and Tomasello 2011), albeit one probably lacking the human ability to represent alternative beliefs or models of the world (Martin and Santos 2016). In contrast, it has been demonstrated that most adult humans have an ability for between four and five orders of intentionality, with an apparent human ceiling of six orders (Kinderman et al. 1998; Stiller and Dunbar 2007). A related concept is that of “common knowledge” (De Freitas et al. 2019) which posits a type of recursive mentalizing, involving, in effect, limitless orders of intentionality, centered on salient focal points, which, it is claimed, facilitates mutual cooperation.

It may be argued, therefore, that although human children have been observed to imitate each other more faithfully than do chimpanzees (Whiten et al. 2009), the Girardian emphasis on the uniqueness of human mimicry may, nevertheless, be misplaced. Mimicry of conspecifics is observed in many species (Zentall 2001). Rather, the trait that perhaps better defines our species may instead be our tendency to attribute intentional states to others, and our ability to do so to multiple orders of intentionality.

To this author's knowledge, Girard did not address the issue of mentalizing, let alone the concept of orders of intentionality, but instead, his theory would appear to take for granted the human ability to so. Yet, in order to mimic the desires of each other, we need first to be able to impute (accurately or otherwise) what those desires might be. It may therefore have been the emergence of such an ability to mentalize that better explains the phenomenon, described by Girard, of the tendency for the desires and other mental states of individuals in a community to spontaneously align. In other words, once a capacity to mentalize emerged, where there may have been a preexisting tendency to imitate, then imitation would no longer have been confined to mimicking each other's actions, but could also have involved mimicking each other's mental states.

## Shared Intentionality

Girard's concept of mimesis is resonant, not only with the idea of common knowledge described above, but also with Michael Tomasello's concept of "shared intentionality" which describes a human tendency to seek, and to create with conspecifics, a shared attentional space, or a joint, or "we" intentionality (Tomasello and Carpenter 2007). As well as facilitating multiple individuals to work towards a joint goal, this also allows for role swapping (Moll and Tomasello 2007). Although employing a very different epistemological method to Girard's (experimental comparative psychology, in contrast to Shakespearian exegesis), there are nevertheless, strong parallels between Tomasello's shared intentionality concept and Girard's mimesis (Kruger 2011). Both describe the spontaneous alignment of intentional states amongst humans, and both consider this to be a species-defining trait. Both are also predicated on a highly developed ability to mentalize and to do so recursively.

Girard emphasizes the process (mimicry of mental states), and Tomasello the outcome (the emergence of a joint attentional framework), yet it is arguable that they are both describing the same phenomenon. However, whereas Tomasello emphasizes the fitness advantages of such a shared attentional framework, especially in terms of cooperation and cumulative culture (Tennie et al. 2009), Girard (being a Shakespearian scholar) focuses instead on the potential rivalries, conflicts, and consequences for social stability.

## Emergence and Stability of Dominance Hierarchies

Social hierarchies are evident in many animals, including primates (Watts 2010). As dominant animals have better access to food and mates, they therefore have direct fitness advantages, hence the emergence of behavioral traits that tend to establish and maintain a dominant position can be readily explained. However, a stable hierarchy is also reliant on the existence, in lower-ranked individuals, of behavioral traits that facilitate submission to dominants. The existence of such "submission traits" might be accounted for both from a direct fitness perspective through the avoidance of serious injury (Maynard Smith and Price 1973), and also in terms of the inclusive fitness advantages (Hamilton 1964; Levin and Grafen 2019) that pertain to a stable group hierarchy (Price 1967). In other words, the existence of both dominance traits and submission traits may be understood as a trade-off between the stress of low rank and the stress of conflict (Debreuil 2010, p. 53).

Price et al (1994) posit that humans share with other animals a "mechanism for yielding in competitive situations," and that hierarchies are the product of "a phylogenetically ancient mechanism for creating social asymmetry between previously equal adults." They describe how a

stable hierarchy could be an emergent property of multiple binary contests whereby, typically, a lower-ranked animal challenges a higher-ranked one. The outcome of such challenges, or "ritualized agonistic behaviors" (Price 1969), might depend on numerous variables (Parker 1974), yet despite the high stakes in terms of direct fitness, typically one or other of the animals concedes, thereby either creating or maintaining social asymmetry.

Maynard Smith (1974), using hawk-dove game theory, demonstrated how a stable distribution of aggressive and submissive traits could become fixed in a population. Such models are simplifications, and complicating factors, amongst primates at least, may be bystander intervention and a tendency to form alliances (Ridgeway and Diekema 1989). Nevertheless, the important point is that conspecific conflict may be curtailed in other animals by the spontaneous emergence of hierarchies that depend for their stability upon a widespread propensity to submit to a dominant. In other words, social stability may rely on most members of a community regularly adopting a different intentional state to that of a conspecific.

The spontaneous adoption of such different intentional states may be the typical outcome of contests amongst species that have little or no ability to infer each other's intentional states, let alone have a propensity to reciprocally imitate them. In humans, however, the tendency to imitate intentional states, or for a shared intentionality to spontaneously emerge, may have been incompatible with maintaining the required differences within a group upon which social hierarchies, and therefore social stability, would have depended.

## Dominance, Cooperation, and Social Stability

Whereas the above discussion, following Girard, considers the effects that the emergence of certain social traits may have had on the stability of dominance hierarchies, other theorists have considered the inverse, that is, the effect that the persistence of dominance hierarchies would have had on the emergence of human sociability, and specifically on cooperativeness (Boehm 1999; Debreuil 2010; Cronin et al. 2015; Tomasello 2016). They suggest that the replacement of a dominance-based social order with a more egalitarian structure would have been a prerequisite for the emergence of cooperation, as participants in a collaborative venture would need to have had confidence that any spoils would be shared equitably.

There is evidence that animal-type dominance hierarchies did indeed become less prominent during human evolution, with early hominins, and even contemporary hunter-gatherers, living in small, apparently egalitarian, acephalous bands, lacking any clear leader or chief, (Knauff et al. 1991; Boehm et al. 1993). Notwithstanding the subsequent reemergence of

human hierarchical societies in more recent millennia, this propensity for egalitarianism has been described as one of the pillars of the early human sociocultural niche (Whiten and Erdal 2012). Tomasello suggests that the emergence of such relative egalitarianism facilitated cooperative activities, which would have allowed early hominins to greatly expand their foraging and hunting opportunities (Tomasello et al. 2012). Indeed, it might be said that the corollary of Girard's comment, concerning mimesis making animal societies impossible, also appears to be true. That is, where a certain threshold of animal-type dominance hierarchy persisted, human (cooperative) societies would have been impossible.

## Self Domestication

An alternative approach to explaining the emergence of conditions compatible with cooperative societies focuses on the propensity for aggression, rather than on dominance or hierarchy, and specifically on the distinction between two types of aggression, reactive and proactive (Dodge and Coie 1987; Wrangham 2018). Reactive aggression is characterized by emotional arousal and impulsivity, whereas proactive aggression tends to be goal directed, and planned. Richard Wrangham (Wrangham 2019b) argues that one of the defining characteristics of our species is the (relatively) low propensity for reactive aggression, and that this is analogous to similar traits that characterize domesticated animals. Such a domestication syndrome (Darwin 1868), so called as it tends to be associated with other incidental traits such as paedomorphism, can, in most cases, be readily accounted for by generations of selection by humans against reactive aggression traits in their animals (Belyaev 1969). However, a domestication syndrome has also seemingly emerged spontaneously in some species, most notably in humans (Wrangham 2019a), but also in one of our closest ape relatives, bonobos (Hare et al. 2012).

Reasoning that species that have undergone self-domestication must have done so in response to environmental pressures that selected against reactive aggression, the execution hypothesis (Hare et al. 2012; Wrangham 2019b) proposes that, in the case of bonobos and humans at least, such selection pressures were mediated through the emergence of coalitions that engaged in proactive aggression against their more reactively aggressive conspecifics. The formation of such aggressive coalitions is evident amongst adult female bonobos who have been observed to make coordinated attacks against overly aggressive males in their own group (Hare et al. 2012). This contrasts with their close relatives, chimpanzees, who typically engage in proactive male coalitions against solitary outgroup members (Mitani et al. 2002), and do not exhibit a domestication syndrome.

Wrangham argues that similar self-domestication by proactive aggressive coalitions occurred, independently, late in the hominin lineage and links this to the emergence of language which would have allowed for greater coordination and conspiratorial planning of such coalitions of proactive aggression (Wrangham 2019a). Reasoning that as cooperation and social learning would have been incompatible with a high propensity to react aggressively during encounters with conspecifics, Wrangham concludes that whereas the direct consequence of proactively aggressive coalitions may have been a domestication syndrome, an incidental consequence, in humans at least, may have been the facilitation of cooperation (Wrangham 2019b, p. 189).

Girard wrote in the mid to late 20th century, in the aftermath of two great European wars and in a world threatened with nuclear war. He takes a view similar to that of Thomas Hobbes (Hobbes and Oakeshott 1946) (who, in the 17th century, also wrote in the aftermath of catastrophic European wars), and regarded humans as a species engaged in a continuous struggle to suppress our self-destructive violent tendencies. Yet, this view of humanity, as a uniquely violent species, appears to be questionable, as our two closest relatives, chimpanzees, and, to a lesser extent, even the more peaceful bonobos, have been observed to kill each other as often as human hunter-gatherers did, while having a much higher frequency, by an order of magnitude, of nonlethal aggressive encounters with each other (Wrangham et al. 2006). Hence, Girard's claim that collective murder resulted in a "true rupture" between animals and humans would seem to be incorrect.

Wrangham presents a more nuanced view on human violence (Wrangham 2019b), arguing that we have a reduced propensity for reactive aggression but a greater propensity for proactive. Girard's hypothesis might be similarly expressed as entailing proactive aggression, in the form of sacrificial ritual, keeping reactive aggression in check, but with the sacrificial rituals themselves having their origin in acts of spontaneous reactive aggression.

## Conclusion

The consensus amongst evolutionary theorists would seem to agree with Girard in considering that some aspects of human sociality would have been incompatible with animal-like dominance hierarchies. However, whereas the focus of some theorists has been on the problems that dominance would have caused for cooperativeness, Girard's emphasis is on the problems that the necessary weakening of dominance systems would have caused for social cohesion. He argued that because the primate dominance hierarchical system had an important role in curtailing intragroup violence, that the transition to a seemingly more egalitarian social structure would have presented significant challenges, and therefore

requires explanation. Like Wrangham, he also posits a role for collective killing, although, as discussed below, Wrangham and Girard differ somewhat in how they envisage both the form of such killings, and the effects they would have had.

### Comparative Analysis; the Posited Solution; the Scapegoat Mechanism

It matters little, in effect, who is the victim, provided that there is one.

(Girard et al. 1978, p. 142)

The main Girardian claim being examined by this article is not so much that traits essential to human sociality would have presented early hominins with a novel fitness challenge in the form of increased conspecific violence, but rather that the scapegoat mechanism emerged as an adaptive response to this challenge.

Humans engage in a spectrum of victimization behaviors (Forgas et al. 2005). These have different names depending on the context, such as ostracization, social exclusion, bullying, or lynching. What these all have in common is a dynamic whereby one individual, or a small group of individuals, is singled out and becomes the target for hostility on the part of the rest of the community. Although analogous behaviors such as “mobbing” are observable in some other animals (Ostreiher 2003), these are invariably directed at intruders, or predators of another species. Humans are unusual in having a propensity to single out for attack, sometimes seemingly at random, not only members of other groups, such as in the case of headhunters (Rosaldo 1980), but also members of their own group, although, as noted above, a similar, albeit seemingly less arbitrary, phenomenon has been observed in both of our closest relatives, bonobos and chimpanzees.

Several other authors have addressed the issue of ostracization (Eisenberger 2003; Forgas et al. 2005; MacDonald and Leary 2005; Sweeting et al. 2006), but invariably do so from the perspective of the serious consequences for the ostracized. Girard in contrast focused on the effect on those doing the ostracizing, claiming that the human propensity for spontaneous victimization provided the framework for generating communal unanimity founded upon a shared sacred meaning. This claim is discussed below in the context of other hypotheses regarding the evolutionary origins of cooperation, morality, religion, ritual, and sacrifice, and on how hierarchies have, or at times seemingly have not, operated during hominin history.

### Other Hypotheses on Religion and Morality

Numerous hypotheses have been advanced regarding the evolutionary origins of the uniquely human phenomenon of religion (Boyer and Bergstrom 2008; Bulbulia et al. 2008). Some of these consider how religious thinking might have emerged out of preexisting cognitive biases that had originally served other functions (Atran and Norenzayan 2004). For instance, it has been proposed that a hypersensitivity towards agency detection that would have been protective against predators, combined with the emergence of an ability to mentalize, would have resulted in a human tendency to ascribe agency to inanimate objects (Atran and Henrich 2010). Similarly, it has been suggested that while a propensity to better remember minimally counterintuitive information, coupled with a tendency to imitate exactly the behaviors of others, might have facilitated intergenerational transmission of skills and knowledge, it would also have allowed religious-type belief systems and associated rituals to become established through cultural selection (Boyd and Richerson 1985; Barrett 2008; Lewis and Laland 2012).

Others have emphasized the emotional (Dunbar 2020), and ritualistic (Rappaport 1999), rather than cognitive, aspects of religion. Dunbar argues that religion is important to the stability of human groups, with evidence that group fission is less, and thus group longevity greater, in communes with a religious rather than a secular ethos (Dunbar and Sosis 2018). Rappaport linked ritual to the emergence of symbolism and language, and argued that archaic religious rituals involved self-referent “indices” which were impervious to falsification, thereby helping to “ameliorate problems of falsification intrinsic to language.” Although this places a greater emphasis on communal practice, rather than on the cognitive aspects of belief, it has similarities with a cognitivist hypothesis that suggests that ritual sacrifices may be understood as credibility enhancing displays (CREDs) which would have facilitated the cultural transmission of belief systems (Henrich 2009).

Although some of these hypotheses disagree on whether religion is better thought of as an adaptation, or as a by-product of other traits (Sosis 2009), there is a consensus that it played a role in communal cohesion and the emergence of group norms and shared values (Durkheim 1915), with some linking it to the relatively large size of typical hominin groups (Dunbar 2008). It has also been proposed that the dramatic Holocene expansion of human community size, which has occurred in recent millennia, was linked to the emergence of a religious cultural variant that involved belief in all powerful deities, so called “Big Gods,” who take a direct interest in human moral behavior (Norenzayan et al. 2016; Lang et al. 2019). Similarly, Dunbar surmises that the

“centerpiece” of the Neolithic revolution wasn't agriculture as such, but rather the emergence of a form of religion that provided “mechanisms that allowed people to live together in cramped conditions without killing each other” (Dunbar 2020).

Another pertinent hypothesis is Tomasello's account of the origins of human morality (Tomasello 2016). This considers the emergence of concepts such as sacred values and norms, and links these to human cooperation. Tomasello has proposed an *interdependence hypothesis* to explain the emergence of such cooperation, beginning with small scale mutual collaboration between individuals founded on a joint intentionality, which evolved into larger-scale group-level collaboration based on “collective intentionality” (Tomasello 2014).

Any upscaling of group size would have been at the cost of reduced individual incentive and reduced information regarding the skills and reputation of partners, thereby increasing the problem of “free riders” and cheaters (Hardin and Garret 2003). Cosmides and Tooby (Cosmides and Tooby 2002) propose that much of human psychology is comprised of dedicated intelligence modules that emerged as specific adaptations to specific problems, and there is evidence for such an adaptation that produces punitive sentiments specifically targeted at free riders (Price et al. 2002).

Cooperative endeavors would have required not only a system for detecting and punishing cheating, but also a shared agreement as to what the behavioral norms were, and how to deal with transgressors (Fehr and Fischbacher 2004; Gardner and West 2004). Tomasello suggests that collective intentionality would have relied on the emergence of new cognitive processes that addressed the problems caused by having to cooperate with relative strangers. This, he claims, would have facilitated the emergence of a collective or cultural common ground, which came to form the basis of shared social norms and what we might call an objective morality (Tomasello 2016, p. 102).

### Comparisons with the Scapegoat Hypothesis

What each of the above approaches have in common is the suggestion that codes of unanimously accepted behavioral norms became established through cultural selection and that this selection was mediated by fitness advantages associated with some form of enhanced group cohesion. The scapegoat hypothesis makes a broadly similar claim. Although it may not explicitly articulate the concept of cultural selection, this is, nonetheless, implicit in Girard's account, as the idea of cultural contagion is a central theme of his work.

What others have described as a “group minded moral psychology” (Tomasello 2016, p. 136), “moral norms” (Atran and Henrich 2010), or “common knowledge” (De Freitas et al. 2019) are, arguably, the same phenomenon as

what Girard described as “the sacred,” as they all involve values or beliefs that are shared by the entire social group and perceived as objectively true. Furthermore, Girard is not alone in linking this phenomenon to either punishment or victimization, with clear parallels between his scapegoat mechanism and Tomasello's hypothesis regarding the punishment of transgressors and the associated emergence of social norms which in turn are linked to group cohesion and meaning. What Tomasello and others describe in terms of a cognitive tendency for third-party norm enforcement and the related creation of cultural institutions that allow punishments to be delivered impersonally or by the group (Dubreuil 2010, p. 189; Tomasello 2016, p. 104), is analogous to what Girard called “sacred violence,” which includes ritual violence, and punishments sanctioned by a deity.

Dunbar argues that early hominins were able to enjoy the advantages of larger group sizes by evolving novel ways of bonding, through the release of endorphins, which were more efficient at creating and maintaining bonds between group members than were the physical contact grooming methods used by other primates (Dunbar 2020). He surmises that the phylogenetically oldest of these uniquely human bonding methods is laughter, which he claims helped to bridge the “bonding gap” between conventional apes and archaic humans (Dunbar 2012). Laughter is a common associate of modern ostracization behaviors (Klages and Wirth 2014) and may also trigger endorphin release (Dunbar et al. 2012). The scenario envisaged by Girard, of spontaneous cathartic victimization, would, therefore, seem consistent with laughter-induced endorphin release playing an important role in early hominin group bonding. Dunbar also points out that the later emergence of a mental capacity to conceptualize four to five orders of intentionality would have been a necessary condition for the type of “storytelling” that is an essential part of communal religions involving social norms (Dunbar 2008). This is analogous to the Girardian concept of the role of mythology, founded on features of the scapegoat mechanism, such as double transference and misrecognition, which would have been similarly reliant on multiple orders of intentionality.

There are, however, some important differences between the scapegoat hypothesis and some of the approaches discussed above. Whereas all accounts envisage the emergence of a process whereby perceived transgressors are punished through some form of ostracization, Girard considered the primary selection pressure to have come from the problem of intragroup conflict, rather than that of free riders. Some other models emphasize the sensitivity of the cheater detection mechanism (Tomasello et al. 2012, p. 679), implying high specificity, with the hypothesized function primarily being the modification of the behavior of accused transgressors. This assumes that the individual accused of norm violation is invariably guilty. Although Girard also envisaged

a high sensitivity to cheater detection, he considered the specificity to be of less relevance, as it is the endorsement of group values that he saw as being of greater functional importance. This is consistent with the argument that “bottom-up” social norms function better at addressing the problem of free riders than do “top-down” transgressor sanctions (Ostrom 1990).

In Girard’s account, the actual guilt or innocence of the accused becomes immaterial because, provided that the group unanimously believe in their guilt, the process of norm enforcement works equally well with a victim who is innocent as with one who is guilty. In other words, accusations primarily function by modifying the beliefs and behaviors of the rest of the group, rather than by any effect on the accused. It is this focus on the role of the *act of accusation*, rather than the *substance of the accusation*, that, perhaps, best illustrates the difference between Girard and most other comparable theorists.

As with the potential, discussed above, for human desire to become divorced from functional utility, the mimetic nature of accusations would have resulted in a similar potential to become detached from any plausible guilt or innocence, resulting in a tendency to accuse individuals of a limitless variety of crimes, some of which they could not possibly have been guilty of, such as in the case of the medieval European phenomenon of witch trials. Such impossible accusations could not have functioned as means of instrumental behavioral modification, but instead may have functioned by attributing to the accused a malign disposition imbued with a mystique likely to generate myth.

A hypothesis that has strong parallels with Girard’s is that of DeScioli and Kurzban (DeScioli and Kurzban 2013), which seeks to account for the emergence of deontological (non-consequential) morality, contrary to the theoretical expectation that evolution should have favored consequentialism. They suggest that deontological thinking (which is arguably analogous to sacred thinking), evolved as a mechanism that facilitated bystanders to coordinate which side to take in a dispute, thereby curtailing what they call “third party discoordination” and the associated escalation of conflicts. Like Girard, they argue that moral cognition functions by coordinating condemnation, rather than by promoting any specific altruistic behaviors or direct beneficial consequences. This, they suggest, allowed the emergence in hominin societies of a delicate balance between despotism (akin to dominance hierarchies) and discoordination (akin to mimetic crises). Girard made similar claims regarding the scapegoat mechanism, but in addition posited that the victim’s ambiguous status, being both demon and deity, villain and king, together with the associated mythology, played an important role in maintaining such a balance.

Regardless of whether it is deontological or consequentialist, the human phenomenon of third-party punishment

has been identified as being of importance to cooperation. Because there is a potential cost for the punisher, this has been described as an altruistic behavior, with debate as to how it may have emerged (Hill et al. 2009). As the greater the proportion of a population who act as punishers, the less the individual-level selection operating against them (Boyd et al. 2003), it has been suggested that, although the maintenance of third-party punishment behavioral traits in a population may be readily accounted for, it is harder to explain how such behaviors originally evolved (Gardner and West 2004). Girard’s concept of spontaneous unanimous hostility generated by mimesis provides a possible solution, as it describes the emergence of what would have been, *ab initio*, widespread, and thus potentially stable, third-party punishment behaviors.

## Human Sacrifice

Human sacrifice, although anathema to modern cultures and religions, appears to have been a widespread practice in pre-modern societies (Bremmer 2007; Acevedo and Thompson 2013) with a geographical distribution suggesting that it was not the result of local cultural aberrations, but rather was the product of a behavioral tendency ubiquitous in humans. Although Girard’s hypothesis does not offer any obvious explanation for some of the more puzzling, but seemingly pervasive, facets of the phenomenon, such as that of child sacrifice (Stager and Wolff 1984; Wilson et al. 2007), it is arguably consistent with other hypotheses that posit an adaptive function for the practice.

The social-control hypothesis suggests that human sacrifice was motivated by political control and intimidation (Winkelman 2014), and this is supported by evidence linking the practice with greater, and more stable, social stratification (Watts et al. 2016). This would seem consistent with Girard’s view that sacrificial rituals played a role in maintaining social differences. Arguably, Girard offers a parsimonious account of how such practices may have initially emerged from spontaneous acts of communal aggression which would, initially at least, have been devoid of any conscious political motivation or goal.

Whether sacrifice functioned by political intimidation or, as suggested by theorists such as Rappaport and Henrich, by reinforcing intergenerational transmission of unanimous communal values, Girard’s concepts of misrecognition and double transference offer an account of how sacred belief systems, symbolism, and language might have become associated with collective killing.

## Execution Hypothesis

There are similarities between Girard’s scapegoat mechanism and the phenomenon of coalitions of aggression

described above, to which Wrangham attributes self-domestication. These coalitions are characterized by the assailants having an overwhelming numerical advantage, thereby taking little personal risk (Wrangham 2019b). Judicial executions are a modern example of such proactive killing of selected, typically single, victims and some form of this seems to be ubiquitous in human societies (Otterbein 1986). Such coalitions necessarily involve individuals having some ability to discern each other's intentions with sufficient accuracy, and the observation of a version of the phenomenon in chimpanzees and bonobos may, arguably, be consistent with it being associated with an ability for at least rudimentary second-order mentalizing (O'Connell and Dunbar 2003; Call and Tomasello 2011).

The execution hypothesis contrasts in several ways with Girard's scapegoat mechanism, perhaps the most significant of which concerns the type of aggression envisaged. Wrangham describes coalitionary aggression as proactive. Indeed, the execution hypothesis essentially posits that, in both humans and bonobos, reactive aggressive traits were selected against by targeted proactive aggression. In contrast, although Girard does not make the distinction between proactive and reactive aggression, he nevertheless clearly envisages the initial scapegoat killing as a collective act of reactive aggression. He argues that lynch mobs do not plan their actions but rather "know not what they do" (Girard 2017, p. 86; English Standard Version Bible 2001; Luke 23:34). However, subsequent reenactments and deliberate human sacrificial rituals would seem to entail proactive aggression and do have some similarities with execution (Smith 2000).

Wrangham and Girard take contrasting views on the primary role played by language. Wrangham focusses on how sophisticated communication would have allowed for complex conspiratorial planning (Wrangham 2019a), whereas Girard, like Dunbar, considers that the more important role played by language related to mythologizing, storytelling, and the intergenerational transmission of sacred values. Arguably, this is a difference in emphasis, with language potentially having played both roles.

They also differ in terms of the hypothesized consequences. Wrangham's execution hypothesis assumes that individuals who exhibited certain behavioral traits were more likely to be selected as victims and that as a result, the genetic basis for such traits was selected against, whereas because Girard regards the identity of the victim to be arbitrary, and of little relevance to the pacifying function, he does not consider the possibility of selection against any particular trait. In other words, Wrangham emphasizes the genetic consequences, Girard the cultural, although these two distinct consequences might be regarded as having been complementary rather than contradictory.

## Hierarchies

As alluded to above, some anthropologists have suggested that hierarchical social structures have had a U-shaped trajectory during human evolution (Knauf et al. 1991; Boehm 1999; Dubreuil 2010), and that in contrast to most apes and to modern humans, earlier hominins and hunter-gatherers appear to have lived in more egalitarian groups. It has been suggested that such egalitarianism was the product of a system of "reverse dominance" (Boehm et al. 1993) which emerged from the tendency of subordinate males in the case of chimpanzees, or females in the case of bonobos, forming alliances against dominant males. Boehm argues that, in hominins, this evolved into a communal intolerance of dominants, or "upstarts," giving rise to what he calls "the domination by the rank and file," resulting in groups acting as "unified moral communities" (Boehm 1999).

There are parallels between this and Girard's scapegoat mechanism such that both may be describing the same development. What appears as egalitarianism may in effect have been a binary hierarchy founded on the difference between the community as a whole and any individual who was either overly dominant (Boehm), or perceived (possibly erroneously) to be transgressing in some way (Girard). Boehm described this as a "very special type" of hierarchy and proposed a definitive attitudinal switch from "don't fool with me" to "don't fool with us." This is resonant with the Girardian idea of a switch from "all against all" to "all against one."

Despite such similarities, an important difference is that Boehm proposes an explicit egalitarian ethos whereby members of a community "deliberately take charge of their own fate" and may even have known what they were doing. In contrast, Girard proposed a more spontaneous and unconscious dynamic, an important part of which would have been the generation of mythology, thereby requiring that the community did *not* understand their own actions.

Social hierarchical structure has reemerged, and did so contemporaneously with the appearance of larger, more complex human societies at the beginning of the Holocene era, about 12,000 years ago. Boehm's account suggests that this might be understood in terms of the dominants eventually regaining the upper hand in their struggle with subordinates. Dubreuil has disputed this interpretation (Dubreuil 2010), arguing instead that modern human hierarchies differ qualitatively from the earlier dominance-based variety in that they are characterized by symbolism and political institutions, which he links to the expansion of the temporoparietal area during the transition from *Homo heidelbergensis* to *Homo sapiens*, some 160,000 years ago, and to the emergence of sufficient attentional flexibility "to look simultaneously at a person as a man or as a president" (Dubreuil

2010, p. 136). Such an ability to hold ambiguous perspectives on individuals is analogous to Girard's account of victims acquiring an ambiguous status of being simultaneously both malign and benign. Thus, Dubreuil and Girard both posit that modern political institutions have their origins in status ambiguity of individuals, although Dubreuil does not ascribe any specific role to victimization in generating this.

## Implications

There are various philosophical, political, and even theological issues raised by Girard's scapegoat hypothesis, and there is a growing body of literature discussing these (Cowdell et al. 2012; Vattimo 2012; Alison et al. 2017). This article does not dwell on these aspects but considers instead some potential implications for evolutionary psychology and psychiatry.

## Protective Traits

If the scapegoat mechanism was an important adaptation in human evolution, then it would have functioned by enhancing the fitness of groups, but with high direct fitness costs for the individuals selected as victims. We might therefore expect to observe in humans both population-wide traits that facilitate the mechanism, but also traits that protect against victimization at the individual level. The risk of being selected as a scapegoat victim would have been, and perhaps still is, an ever-present threat in the human social environment. Hence the sensitivity of the human psyche to the risk of social exclusion, as evidenced by our susceptibility to conditions such as social anxiety and social phobia, may reflect not only the intense social nature of the human niche and the adverse implications of reputational damage, but may also be understood in terms of the potentially high direct fitness costs associated with being singled out from the crowd.

## Functional Victims

Traits that function by promoting individual self-preservation may be readily understandable. However, if scapegoating were adaptive, and its function mediated not through the actual guilt or innocence of the victims, but rather through how they were perceived by others, then the prospect arises of traits having also emerged, in a small number of individuals, which rather than functioning through the avoidance of victimization, would instead have attracted it. In

other words, there may have been selection pressures for the emergence of individuals with phenotypes that would have made them more efficacious scapegoats. Such "victimization traits" would have functioned as a form of altruism, increasing inclusive fitness by enhancing the efficacy of the scapegoat mechanism.

This concept of victim phenotype is the basis of the *archetypal scapegoat hypothesis* (ASH) on the evolutionary origins of psychosis (Riordan 2017). This posits that psychosis emerged as an adaptive response to the risks of mimetic internecine violence by providing groups with individuals who were more likely to engender unanimous hostility from others, and who were already imbued with a mystique that facilitated double transference. The latter function, that of psychosis having a role in enhancing communal unity by inducing religious thinking, has been proposed by a number of previous evolutionary hypotheses (Polimeni and Reis 2002; Stevens and Price 2015). However, the Girardian perspective of the ASH also focuses on the hostility, and unanimity, of the rest of the population. Unanimity would have been of critical importance to the pacifying function of scapegoating, as without it the mechanism would have been "worse than useless" (Girard 1972, p. 113). The ASH suggests that the evolutionary origins of psychosis cannot be understood by studying the individual sufferers alone, but of equal importance is a tendency in the general population to scapegoat the mentally ill (Riordan 2019), a tendency that we now recognize as stigmatization.

This hypothesis is consistent with several known risk factors for psychosis, such as exposure to "high expressed emotions" or trauma (Bechdolf et al. 2010), which might be understandable as environmental associates of mimetic crises, which induce the pacifying effects of the scapegoat mechanism by providing a suitable victim in response to social disharmony. Other risk factors, such as having experienced bullying (Trotta et al. 2013) or being a member of a minority group (Morgan et al. 2019), might be understood in terms of victimization being self-reinforcing, whereby overt psychosis is more likely to develop in a person who has already been singled out, and is already being victimized. The importance of unanimity suggests that frequency-dependent selection may help to explain the relatively uniform prevalence of psychosis, as a higher prevalence would dilute the unanimity-inducing function. Furthermore, the ASH suggests that the apparent changes in societal attitudes to the mentally ill over recent centuries (Foucault 1965) are consistent with the contemporaneous shift in attitudes towards sacrificial religion and victimization in general (Girard 1986; Riordan 2020).

## Conclusion

Up to now it was recognised that the rapid increase in brain size, as well as other phenomena, required an interaction between biological and cultural factors, but what has been missing is a model for the organizing and driving factor in the process, a motor car for this strange machine. The victimage process gives us this motor.

(Girard et al. 1978, p. 96)

This article considers the claim that the scapegoat mechanism, as described by René Girard, played an essential role in the transition of early hominins from hierarchical societies based on dominance to the seemingly more egalitarian social structure required for cooperation. Although Girard's working assumption regarding lethal conspecific violence being unusual in other animals may have subsequently been found to be inaccurate, his hypothesis is nevertheless consistent with other approaches that regard religious or deontological thinking, together with shared communal values, as being central to the emergence of cooperation and the stability of large groups. It also has commonalities with hypotheses that posit a role for aggressive coalitions. What Girard adds, which other hypotheses lack, is the claim regarding the role played by arbitrary victimization and attribution of guilt in generating religious thinking and sacred values. This is, arguably, a parsimonious addition as it posits a mechanism that would have been, for the most part, a spontaneous emergent property of behavioral and cognitive processes, in particular imitation and mentalizing, which are already well recognized, whilst potentially offering an explanatory framework for phenomena as varied as third-party punishment, bullying, witch hunts, human sacrifice, funerary practices, and psychosis.

Whiten and Erdal (2012) proposed five pillars of what they call the "deep social mind," namely egalitarianism, mentalizing, cooperation, language, and cultural transmission. They suggest that these form "an interrelated, adaptive socio-cognitive complex," and propose that this complex evolved by way of two-way positive feedback interactions between these five "pillars." This article proposes that scapegoating should be regarded as a sixth such pillar, as essential to the emergence of the socio-cognitive complex as any of the other five.

René Girard may not have been a trained scientist, psychologist, anthropologist, nor clinician. Indeed, he might more accurately be described as an "ivory tower" academic, with a Freud-like interest in mythology, who took his primary inspiration from the likes of Dostoevsky, and Shakespeare. As such, he may seem an unlikely person of interest for evolutionary psychology. Yet from this cross-disciplinary perspective, his concept of the scapegoat mechanism, and

the role played by victimization in the emergence of humanity, may be a vital missing piece of the jigsaw puzzle of human evolution and the origins of our "strange machine."

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Acevedo GA, Thompson M (2013) Blood, war, and ritual: religious ecology, 'strong' culture, and human sacrifice in the premodern world. *Anthropol Forum* 23(3):266–288. <https://doi.org/10.1080/00664677.2013.805121>
- Alison J, Palaver W, Merrill TC, Hidden ST (eds) (2017) *The Palgrave handbook of mimetic theory and religion*. Palgrave Macmillan, New York. <https://doi.org/10.1057/978-1-137-53825-3>
- Antonello P, Gifford P (eds) (2015) *How we became human: mimetic theory and the science of evolutionary origins*. Michigan State University Press, East Lansing. <https://muse.jhu.edu/book/41704>
- Atran S, Henrich J (2010) The evolution of religion: how cognitive by-products, adaptive learning heuristics, ritual displays, and group competition generate deep commitments to prosocial religions. *Biol Theor* 5(1):18–30. [https://doi.org/10.1162/BIOT\\_a\\_00018](https://doi.org/10.1162/BIOT_a_00018)
- Atran S, Norenzayan A (2004) Religion's evolutionary landscape: counterintuition, commitment, compassion, communion. *Behav Brain Sci* 27(6):713–730. <https://doi.org/10.1017/S0140525X04000172>
- Baron-Cohen S (1995) *Mindblindness: an essay on autism and theory of mind*. MIT Press, Cambridge
- Barrett J (2008) Coding and quantifying counterintuitiveness in religious concepts: theoretical and methodological reflections. *Method Theor Study Religion* 20(4):308–338. <https://doi.org/10.1163/157006808X371806>
- Bateson G (1935) 199. culture contact and schismogenesis. *Man* 35:178–183. <https://doi.org/10.2307/2789408>
- Bechdolf A, Thompson A, Nelson B, Cotton S, Simmons MB, Amminger GP et al (2010) Experience of trauma and conversion to psychosis in an ultra-high-risk (prodromal) group. *Acta Psychiatr Scand* 121(5):377–384. <https://doi.org/10.1111/j.1600-0447.2010.01542.x>
- Belyaev DK (1969) Domestication of animals. *Science* 5(1):47–52
- Boehm C (1999) *Hierarchy in the forest: the evolution of egalitarian behavior*. Harvard University Press, Cambridge
- Boehm C, Barclay HB, Dentan RK, Dupre MC, Hill JD, Kent S et al (1993) Egalitarian behavior and reverse dominance hierarchy [and comments and reply]. *Curr Anthropol* 34(3):227–254. <https://doi.org/10.1086/204166>
- Boyd R, Richerson PJ (1985) *Culture and the evolutionary process*. University of Chicago Press, Chicago
- Boyd R, Gintis H, Bowles S, Richerson PJ (2003) The evolution of altruistic punishment. *Proceed Natl Acad Sci USA* 100(6):3531–3535. <https://doi.org/10.1073/pnas.0630443100>

- Boyer P, Bergstrom B (2008) Evolutionary perspectives on religion. *Annu Rev Anthropol* 37:111–130. <https://doi.org/10.1146/annur.ev.anthro.37.081407.085201>
- Bremmer JN (ed) (2007) *The strange world of human sacrifice*, vol 1. Peeters, Leuven
- Bulbulia J, Sosis R, Harris E, Genet R, Genet C, Wyman K (2008) *The evolution of religion*. Collins Foundation Press, Santa Margarita
- Call J, Tomasello M (2011) Does the chimpanzee have a theory of mind? 30 years later. In: Schleiden S, Jungert M, Bauer R, Sandow V (eds) *Human nature and self design*. Mentis, Paderborn, pp 83–96. [https://doi.org/10.30965/9783957438843\\_008](https://doi.org/10.30965/9783957438843_008)
- Connell O, Dunbar RIM (2003) A test for comprehension of false belief in chimpanzees. *Evol Cognit* 9(2):131–140
- Cosmides L, Tooby J (2002) Unraveling the enigma of human intelligence: evolutionary psychology and the multimodal mind. In: Sternberg RJ, Kaufman JC (eds) *The evolution of intelligence*. Lawrence Erlbaum Associates, Mahwah, pp 145–198
- Cowdell S, Fleming C, Hodge J (eds) (2012) *Violence, desire, and the sacred: Girard's mimetic theory across the disciplines*, vol 1. A&C Black, New York
- Cronin KA, Acheson DJ, Hernández P, Sánchez A (2015) Hierarchy is detrimental for human cooperation. *Sci Rep* 5:18634. <https://doi.org/10.1038/srep18634>
- Darwin C (1868) *The variation of animals and plants under domestication*. John Murray, London
- De Freitas J, Thomas K, DeScioli P, Pinker S (2019) Common knowledge, coordination, and strategic mentalizing in human social life. *Proc Natl Acad Sci USA* 116(28):13751–13758. <https://doi.org/10.1073/pnas.1905518116>
- Dennett DC (1989) *The intentional stance*. MIT Press, Cambridge
- DeScioli P, Kurzban R (2013) A solution to the mysteries of morality. *Psychol Bull* 139(2):477. <https://doi.org/10.1037/a0029065>
- Dodge KA, Coie JD (1987) Social-information-processing factors in reactive and proactive aggression in children's peer groups. *J Pers Soc Psychol* 53:1146–1158
- Dubreuil B (2010) *Human evolution and the origins of hierarchies: the state of nature*. Cambridge University Press, Cambridge
- Dunbar R (1998) The social brain hypothesis. *Evol Anthropol: Issues News Rev: Issues, News, Rev* 6(5):178–190
- Dunbar R (2008) Mind the gap; or why humans are not just great apes. *Proc Br Acad* 154:403–423
- Dunbar R (2012) Bridging the bonding gap: the transition from primates to humans. *Philosophic Trans Royal Soc b: Biol Sci* 367(1597):1837–1846. <https://doi.org/10.1098/rstb.2011.0217>
- Dunbar R (2020) Religion, the social brain and the mystical stance. *Arch Psychol Relig* 42(1): 46–62. <https://doi.org/10.1177/2F0084672419900547>
- Dunbar R, Sosis R (2018) Optimising human community sizes. *Evol Hum Behav* 39(1):106–111. <https://doi.org/10.1016/j.evolhumbehav.2017.11.001>
- Dunbar R, Baron R, Frangou A, Pearce E, Van Leeuwen EJ, Stow J et al (2012) Social laughter is correlated with an elevated pain threshold. *Proc Royal Soc b: Biol Sci* 279(1731):1161–1167. <https://doi.org/10.1098/rspb.2011.1373>
- Durkheim E (1915) *The elementary forms of the religious life* (trans: Swain JW). George Allen & Unwin, London
- Eisenberger NI, Lieberman MD, Williams KD (2003) Does rejection hurt? An fMRI study of social exclusion. *Science* 302(5643):290–292. <https://doi.org/10.1126/science.1089134>
- English Standard Version Bible (2001) <https://www.biblegateway.com/versions/New-English-Translation-NET-Bible/>
- Evans-Pritchard EE (1962) *Social anthropology and other essays: combining social anthropology and essays in social anthropology*. Free Press, New York
- Fehr E, Fischbacher U (2004) Third-party punishment and social norms. *Evol Hum Behav* 25(2):63–87. [https://doi.org/10.1016/S1090-5138\(04\)00005-4](https://doi.org/10.1016/S1090-5138(04)00005-4)
- Fleming C (2004) *René Girard; violence and mimesis*. Polity Press, Cambridge
- Forgas J, Williams KD, von Hippel W (eds) (2005) *The social outcast: ostracism, social exclusion, rejection, and bullying*. Psychology Press, New York/Hove
- Foucault M (1965) *Madness and civilization* (trans: Howard R). Random House, New York
- Frazer JG (1913) *The golden bough*, vol. 9, the scapegoat. Macmillan and Company, London.
- Gardner A, West SA (2004) Cooperation and punishment, especially in humans. *Am Nat* 164(6):753–764. <https://doi.org/10.1086/425623>
- Girard R (1961) *Mensonge romantique et verité romanesque*. Editions Bernard Grasset, Paris. Girard R (1976) *Deceit, desire, and the novel: self and other in literary structure* (trans: Freccero Y). Johns Hopkins University Press, Baltimore
- Girard R (1972) *La Violence et le sacre*. Editions Bernard Grasset, Paris. Girard R (1977) *Violence and the sacred* (trans: Gregory P). Johns Hopkins University Press, Baltimore
- Girard R (1986) *The scapegoat*. Johns Hopkins University Press, Baltimore
- Girard R (2017) *Evolution and conversion: Dialogues on the origins of culture*. Bloomsbury Publishing, New York
- Girard R, Oughourlian JM, Lefort G (1978) *Des choses cachees depuis la foundation du monde*. Editions Grasset et Fasquelle, Paris. Girard R (1987) *Things hidden since the foundation of the world* (trans: Bann S, Metteer M). Continuum, London
- Hamilton WD (1964) The genetical evolution of social behaviour II. *J Theor Biol* 7(1):17–52. [https://doi.org/10.1016/0022-5193\(64\)90039-6](https://doi.org/10.1016/0022-5193(64)90039-6)
- Hardin R, Garrett C (2003) The free rider problem. In: Zalta EN (ed) *The Stanford encyclopedia of philosophy* (winter 2020 edn). <https://plato.stanford.edu/archives/win2020/entries/free-rider>
- Hare B, Wobber V, Wrangham R (2012) The self-domestication hypothesis: evolution of bonobo psychology is due to selection against aggression. *Anim Behav* 83(3):573–585. <https://doi.org/10.1016/j.anbehav.2011.12.007>
- Haven CL (2018) *Evolution of desire: a life of René Girard*. Michigan State University Press, East Lansing
- Henrich J (2009) The evolution of costly displays, cooperation and religion: credibility enhancing displays and their implications for cultural evolution. *Evol Hum Behav* 30(4):244–260. <https://doi.org/10.1016/j.evolhumbehav.2009.03.005>
- Hill K, Barton M, Hurtado AM (2009) The emergence of human uniqueness: characters underlying behavioral modernity. *Evol Anthropol: Issues News Rev: Issues, News, Rev* 18(5):187–200. <https://doi.org/10.1002/evan.20224>
- Hobbes T, Oakeshott M (1946) *Leviathan: or, the matter, forme and power of a commonwealth, ecclesiasticall and civil*; edit. with a introd. by Michael Oakeshott. Blackwell, Oxford
- Kinderman P, Dunbar R, Bentall RP (1998) Theory-of-mind deficits and causal attributions. *Br J Psychol* 89(2):191–204. <https://doi.org/10.1111/j.2044-8295.1998.tb02680.x>
- Klages SV, Wirth JH (2014) Excluded by laughter: laughing until it hurts someone else. *J Soc Psychol* 154(1):8–13
- Knauff BM, Abler TS, Betzig L, Boehm C, Dantan RK, Kiefer TM et al (1991) Violence and sociality in human evolution [and comments and replies]. *Current Anthropol* 32:391–428. <http://ringmar.net/mycourses/wp-content/uploads/2017/10/Knauff-et-al-1.pdf>
- Kruger AC (2011) Imitation, communion, and culture. In: Garrels S (ed) *Mimesis and science: empirical research on imitation*

- and the mimetic theory of culture and religion. Michigan State University Press, East Lansing, pp 111–127
- Landy J (2012) Deceit, desire, and the literature professor: why Girardians exist. *Republ Lett* 3(1):1–21
- Lang M, Purzycki BG, Apicella CL, Atkinson QD, Bolyanatz A, Cohen E et al (2019) Moralizing gods, impartiality and religious parochialism across 15 societies. *Proc R Soc B* 286(1898):20190202. <https://doi.org/10.1098/rspb.2019.0202>
- Levin SR, Grafen A (2019) Inclusive fitness is an indispensable approximation for understanding organismal design. *Evolution* 73(6):1066–1076. <https://doi.org/10.1111/evo.13739>
- Lewis HM, Laland KN (2012) Transmission fidelity is the key to the build-up of cumulative culture. *Philosophic Trans Royal Soc B: Biol Sci* 367(1599):2171–2180. <https://doi.org/10.1098/rstb.2012.0119>
- MacDonald G, Leary MR (2005) Why does social exclusion hurt? The relationship between social and physical pain. *Psychological Bull* 131(2): 202
- Martin A, Santos LR (2016) What cognitive representations support primate theory of mind? *Trends Cogn Sci* 20(5):375–382. <https://doi.org/10.1016/j.tics.2016.03.005>
- Maynard Smith J (1974) The theory of games and the evolution of animal conflicts. *J Theor Biol* 47(1):209–221. [https://doi.org/10.1016/0022-5193\(74\)90110-6](https://doi.org/10.1016/0022-5193(74)90110-6)
- Maynard Smith J, Price GR (1973) The logic of animal conflict. *Nature* 246(5427):15–18. <https://doi.org/10.1038/246015a0>
- Merrill TC (2017) Critiques of Girard's mimetic theory. In: Alison J, Palaver W, Merrill TC, Hidden ST (eds) *The Palgrave handbook of mimetic theory and religion*. Palgrave Macmillan, New York, pp 455–461
- Mitani JC, Watts DP, Muller MN (2002) Recent developments in the study of wild chimpanzee behavior. *Evolut Anthropol: Issues, News, and Rev: Issues, News, Rev* 11(1):9–25. <https://doi.org/10.1002/evan.10008>
- Moll H, Tomasello M (2007) Cooperation and human cognition: the Vygotskian intelligence hypothesis. *Philosophic Transact Royal Soc b: Biol Sci* 362(1480):639–648. <https://doi.org/10.1098/rstb.2006.2000>
- Morgan C, Knowles G, Hutchinson G (2019) Migration, ethnicity and psychoses: evidence, models and future directions. *World Psychiatry* 18(3):247–258. <https://doi.org/10.1002/wps.20655>
- Norenzayan A, Shariff AF, Gervais WM, Willard AK, McNamara RA, Slingerland E, Henrich J (2016) The cultural evolution of prosocial religions. *Behav Brain Sci*. <https://doi.org/10.1017/S0140525X14001356>
- Ostreiher R (2003) Is mobbing altruistic or selfish behaviour? *Anim Behav* 66(1):145–149. <https://doi.org/10.1006/anbe.2003.2165>
- Ostrom E (1990) *Governing the commons: the evolution of institutions for collective action*. Cambridge University Press, Cambridge
- Otterbein KF (1986) *The ultimate coercive sanction: A cross-cultural study of capital punishment*. Hraf Press, New Haven
- Oughourlian JM (2016) *The mimetic brain*. Michigan State University Press, East Lansing
- Parker GA (1974) Assessment strategy and the evolution of fighting behaviour. *J Theor Biol* 47(1):223–243. [https://doi.org/10.1016/0022-5193\(74\)90111-8](https://doi.org/10.1016/0022-5193(74)90111-8)
- Polimeni J, Reiss JP (2002) How shamanism and group selection may reveal the origins of schizophrenia. *Med Hypotheses* 58(3):244–248. <https://doi.org/10.1054/mehy.2001.1504>
- Premack D, Woodruff G (1978) Does the chimpanzee have a theory of mind? *Behav Brain Sci* 1(4):515–526. <https://doi.org/10.1017/S0140525X00076512>
- Price J (1967) The dominance hierarchy and the evolution of mental illness. *Lancet* 290(7509):243–246. [https://doi.org/10.1016/S0140-6736\(67\)92306-9](https://doi.org/10.1016/S0140-6736(67)92306-9)
- Price J (1969) The ritualization of agonistic behaviour as a determinant of variation along the neuroticism/stability dimension of personality. *Proc R Soc Med* 62(11 Pt 1):1107
- Price J, Sloman L, Gardner R, Gilbert P, Rohde P (1994) The social competition hypothesis of depression. *Br J Psychiatry* 164(3):309–315. <https://doi.org/10.1192/bjp.164.3.309>
- Price ME, Cosmides L, Tooby J (2002) Punitive sentiment as an anti-free rider psychological device. *Evol Hum Behav* 23(3):203–231. [https://doi.org/10.1016/S1090-5138\(01\)00093-9](https://doi.org/10.1016/S1090-5138(01)00093-9)
- Rappaport RA (1999) *Ritual and religion in the making of humanity*. Cambridge University Press, Cambridge
- Ridgeway C, Diekema D (1989) Dominance and collective hierarchy formation in male and female task groups. *Am Sociol Rev* <https://doi.org/10.2307/2095663>
- Riordan DV (2017) Mimetic theory and the evolutionary paradox of schizophrenia: the archetypal scapegoat hypothesis. *Med Hypotheses* 108:101–107. <https://doi.org/10.1016/j.mehy.2017.08.010>
- Riordan DV (2019) Scapegoating mentally ill people. *Br J Psychiatry* 215(2):504–505. <https://doi.org/10.1192/bjp.2019.148>
- Riordan DV (2020) The stones the builders rejected: the scapegoat mechanism and evolutionary psychiatry. *J Viol Mimesis Culture* 27: 59–80. <https://doi.org/10.14321/contagion.27.2020.0059>
- Rosaldo R (1980) *Ilongot headhunting, 1883–1974: A study in society and history*. Stanford University Press, Stanford
- Serres M (2009) Receiving René Girard into the Académie Française. In: Goodhart S, Jorgensen J, Ryba T, Williams J (eds) *For René Girard: essays in friendship and in truth*. Michigan State University Press, East Lansing, pp 1–18
- Smith BK (2000) Capital punishment and human sacrifice. *J Am Acad Relig* 68(1): 3–25. <https://www.jstor.org/stable/1465708>
- Sosis R (2009) The adaptationist-byproduct debate on the evolution of religion: five misunderstandings of the adaptationist program. *J Cogn Cult* 9(3–4):315–332. <https://doi.org/10.1163/156770909X12518536414411>
- Stager LE, Wolff SR (1984) Child sacrifice at Carthage—religious rite or population control. *Biblic Archaeol Rev* 10(1):30–51
- Stevens A, Price J (2015) *Evolutionary psychiatry: A new beginning*. Routledge, Abingdon/New York
- Stiller J, Dunbar RI (2007) Perspective-taking and memory capacity predict social network size. *Soc Netw* 29(1):93–104. <https://doi.org/10.1016/j.socnet.2006.04.001>
- Sweeting H, Young R, West P, Der G (2006) Peer victimization and depression in early–mid adolescence: a longitudinal study. *Br J Educ Psychol* 76(3):577–594. <https://doi.org/10.1348/000709905X49890>
- Tennie C, Call J, Tomasello M (2009) Ratcheting up the ratchet: on the evolution of cumulative culture. *Philosophic Trans Royal Soc b: Biol Sci* 364(1528):2405–2415. <https://doi.org/10.1098/rstb.2009.0052>
- Tomasello M (2014) *A natural history of human thinking*. Harvard University Press, Cambridge
- Tomasello M (2016) *A natural history of human morality*. Harvard University Press, Cambridge
- Tomasello M, Carpenter M (2007) Shared intentionality. *Dev Sci* 10(1):121–125. <https://doi.org/10.1111/j.1467-7687.2007.00573.x>
- Tomasello M, Melis AP, Tennie C, Wyman E, Herrmann E (2012) Two key steps in the evolution of human cooperation: the interdependence hypothesis. *Curr Anthropol* 53(6):673–692. <https://doi.org/10.1086/668207>
- Traube E (1979) Incest and mythology: anthropological and Girardian perspectives. *Berkshire Rev* 14:37–54
- Trotta A, Di Forti M, Mondelli V, Dazzan P, Pariante C, David A et al (2013) Prevalence of bullying victimisation amongst first-episode

- psychosis patients and unaffected controls. *Schizophr Res* 150(1):169–175. <https://doi.org/10.1016/j.schres.2013.07.001>
- Vattimo R (2012) Heidegger and Girard: kénosis and the end of metaphysics. In: Antonello P (ed) *Christianity, truth, and weakening faith: a dialogue*. Columbia University Press, New York, pp 78–87
- Watts, DP (2010) Dominance, power, and politics in nonhuman and human primates. In: Kappeler P, Silk J (eds) *Mind the gap*. Springer, Berlin/Heidelberg, pp 109–138. [https://doi.org/10.1007/978-3-642-02725-3\\_5](https://doi.org/10.1007/978-3-642-02725-3_5)
- Watts J, Sheehan O, Atkinson QD, Bulbulia J, Gray RD (2016) Ritual human sacrifice promoted and sustained the evolution of stratified societies. *Nature* 532(7598):228–231. <https://doi.org/10.1038/nature17159>
- Whiten A, Erdal D (2012) The human socio-cognitive niche and its evolutionary origins. *Philosophic Trans Royal Soc b: Biol Sci* 367(1599):2119–2129. <https://doi.org/10.1098/rstb.2012.0114>
- Whiten A, McGuigan N, Marshall-Pescini S, Hopper LM (2009) Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophic Trans Royal Soc b: Biol Sci* 364(1528):2417–2428. <https://doi.org/10.1098/rstb.2009.0069>
- Wilson AS, Taylor T, Ceruti MC, Chavez JA, Reinhard J, Grimes V et al (2007) Stable isotope and DNA evidence for ritual sequences in Inca child sacrifice. *Proc Natl Acad Sci USA* 104(42):16456–16461. <https://doi.org/10.1073/pnas.0704276104>
- Winkelman M (2014) Political and demographic-ecological determinants of institutionalised human sacrifice. *Anthropol Forum* 24(1):47–70
- Wrangham RW (2018) Two types of aggression in human evolution. *Proc Natl Acad Sci USA* 115(2):245–253. <https://doi.org/10.1073/pnas.1713611115>
- Wrangham R (2019) *The goodness paradox: how evolution made us both more and less violent*. Profile Books, London
- Wrangham RW (2019) Hypotheses for the evolution of reduced reactive aggression in the context of human self-domestication. *Front Psychol* 10:1914. <https://doi.org/10.3389/fpsyg.2019.01914>
- Wrangham RW, Wilson ML, Muller MN (2006) Comparative rates of violence in chimpanzees and humans. *Primates J Primatols* 47(1):14–26. <https://doi.org/10.1007/s10329-005-0140-1>
- Zentall TR (2001) Imitation in animals: evidence, function, and mechanisms. *Cybern Syst* 32(1–2):53–96. <https://doi.org/10.1080/019697201300001812>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.