Does Religion Promote Trust?  
The Role of Signaling, Reputation, and Punishment  

Richard Sosis  
Senior Lecturer, Hebrew University of Jerusalem  
and  
Associate Professor of Anthropology, University of Connecticut*  

*richard.sosis@uconn.edu
Does Religion Promote Trust?
The Role of Signaling, Reputation, and Punishment*

Richard Sosis
Hebrew University of Jerusalem and University of Connecticut

Abstract

Scholars of religion, including such luminaries as Durkheim, Rappaport, Turner, and Weber, have widely assumed that religion promotes intragroup trust among adherents. Recent applications of signaling theory to religious behavior among economists, cognitive scientists, and evolutionary anthropologists further endorse this assumption. However, trust has not been rigorously or consistently defined across authors, making generalizations difficult to evaluate. Here I follow Bacharach and Gambetta’s (2001) behavioral definition of trust and show that the conditions for intragroup trust are often not met in religious communities, especially isolationist and closed communities to which high levels of trust are typically ascribed. Rather, in such communities, cooperation is maintained through institutional structures that effectively punish cheaters and enhance the value of an honest reputation. These groups gainfully facilitate collective action by offering a circumscribed social arena in which reputations can be built, evaluated, rewarded, and efficiently punished. While face-to-face reciprocal relations obviate the need for trusting behavior within closed religious communities, when social groups are fluid, religious practices and symbolic markers are successful in promoting trust among in-group members and anonymous coreligionists who reside in different communities. In addition, these religious badges of identity may be used by non-group members as signals of trustworthiness.

*I thank Candace Alcorta, Joseph Bulbulia, Sharon Feldstein, Bradley Ruffle, and two anonymous reviewers for very helpful comments on earlier drafts of this manuscript. I additionally thank the Russell Sage Foundation (Initiative on Trust) and the University of Connecticut for generous funding of this project.
Trust none; for oaths are straws, men’s faiths are wafer-cakes …

William Shakespeare

*The Life of King Henry the Fifth* (Pistol at II, iii)

[A] group within which there is extensive trustworthiness and extensive trust is able to accomplish much more than a comparable group without that trustworthiness and trust.

Coleman (1988: S101)

While it is broadly recognized that both religion and trust are directly relevant to current world affairs, surprisingly little attention has been given to how they are related to each other. This is particularly notable because recent theoretical developments in the study of religion posit that trust plays an important role in generating the cooperatively derived benefits that religious groups are presumed to offer (see Sosis and Alcorta 2003). Economists (Berman 2000; Carr and Landa 1983; Iannaccone 1992, 1994), cognitive scientists (Atran 2002; Bulbulia 2004a, 2004b), and evolutionary anthropologists (Cronk 1994; Irons 1996a, 1996b, 1996c; 2001, 2004; Sosis 2003, 2004; Sosis and Alcorta 2003) have converged in applying signaling theory to model and explain religious behaviors that are puzzling from the perspectives of the egoistic-based models that serve as these disciplines’ foundations (e.g., rational choice, evolutionary game theory). These researchers have argued that religious behaviors signal group commitments, thus enhancing intragroup trust and facilitating collective action.

Various authors working outside the signaling framework have also asserted that religion enhances trust among community members (e.g., Boudon 1987; Collins 2004; Shield 2002; Steadman and Palmer 1995; Weber 1958: 302–322). However, inconsistency in definitions of trust have limited the ability to make generalizations about this relationship. Even in our everyday use of the term, there is ambiguity. Colloquially, we may employ the word *trust* to characterize the attitude of agents toward other agents. Thus trust sometimes denotes an attitude of confidence about another agent’s reliability. However, we also use *trust* to describe behavior. Accordingly, to trust is to act on this attitude of confidence about another agent’s reliability. The statement “Richard trusts Joseph” may refer to a specific attitude that Richard maintains about Joseph, or it may refer to an actual behavior that manifests itself as a consequence of the attitude—for instance, that Richard trusts Joseph with his wallet. In the discussion that follows, all references to *trust* refer to its behavioral usage, not its attitudinal one. This focus on the behavioral aspects of trust is motivated by the fact that observers of religion have generally commented on the remarkable levels of trusting behavior...

---

1 In my trusty *American College Dictionary*, *trust* is defined as a noun, a verb, and an adjective, and 23 definitions are offered. No wonder this term has caused so much confusion!
exhibited by members of certain religious communities,\(^2\) with less emphasis on the more difficult to observe and measure attitudes and values that encourage these behavioral patterns.

To address the question posed in the title of this article, I first offer a definition of trust that can be easily incorporated into signaling models of religious behavior and delineates between various types of trust. I clarify the benefits that enhanced trust has for individuals and groups and the conditions under which we would expect religious communities to exhibit high levels of intragroup trust. I then examine the prevalence of trust within several well-studied religious communities. To summarize the main argument, while many researchers have maintained that religion increases trust, I argue that this may in fact be illusory. Most of what observers claim is trust in religious communities, especially those communities that demand exclusive commitment from their members, is actually a consequence of institutional frameworks that effectively punish cheaters and support gains for those who maintain an honest reputation. These communities, which I refer to as “closed” (see below), can efficiently punish untrustworthiness and offer reputational histories because social relations are bounded. This circumscribed social arena is created by religious rituals and symbolic markers (e.g., clothing, hairstyles) that clearly signal and demarcate in- and out-group membership. However, when group memberships are fluid and individuals maintain multiple group identities, punishment and reputational mechanisms are less influential. Consequently, these “open” religious communities often rely on religiously endorsed practices and symbolic markers to foster trust. Religious behaviors and badges also promote trust between anonymous in-group members residing in different communities and generate epiphenomenal benefits from increased perceived trustworthiness by nongroup members.

**WHAT IS TRUST?**

As any reader of this journal will be aware, religion has as many definitions as there are scholars who have attempted to define it (for reviews, see Klass 1995; Spiro 1966). Trust, on the other hand, does not seem to suffer from this gratuitous attention, as most researchers discuss trust without any clarification about what it actually is. Indeed, while trust is the core issue of all signaling-based theories of religion, none of this literature attempts to define or measure trust. Here, I begin with a definition of trust offered by Bacharach and Gambetta (2001: 151): “a

---

\(^2\) Throughout this article, I use the term *religious community* to denote a self-defined group that is united by a core set of religious beliefs and distinguishing ritual practices. Typically, but not always, religious communities are organized around one or more churches, mosques, synagogues, temples, or some other institutional or noninstitutional houses of worship.
person trusts someone to do X if she acts on the expectation that he will do X when two conditions obtain: Both know that if he fails to do X, she would have done better to act otherwise, and her acting in the way she does gives him a selfish reason not to do X.\(^3\) Bacharach and Gambetta define a simple two-person trust game according to the following payoffs, where \(x < 3\) and \(y > -3\):

<table>
<thead>
<tr>
<th></th>
<th>Trustee</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>B'</td>
</tr>
<tr>
<td>Truster</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>3, 1</td>
<td>-3, 4</td>
</tr>
<tr>
<td>A'</td>
<td>(x, -)</td>
<td>(y, -)</td>
</tr>
</tbody>
</table>

In this trust game, \(A\) and \(B\) are trusting and trustworthy decisions by the Truster and Trustee, respectively, whereas \(A'\) and \(B'\) are distrustful and defecting decisions.

To illustrate, consider the following anecdote about a tourist who visited New York City for the first time. After arriving at Port Authority by bus, he was approached by a man who asked him whether he needed a cab. The tourist was pleased to be offered assistance and consented. The man took both his bags and told him to wait while he pulled his cab around. Bacharach and Gambetta’s definition of trust accurately characterizes the payoffs in this interaction. The Truster stood to gain an easy cab ride to his hotel but risked losing his luggage. The Trustee stood to gain a cab fare or, more profitably, two pieces of luggage and their contents. Unfortunately for this tourist, who happens to be a friend of mine, the Trustee chose the latter, and the Truster never saw his belongings again.

While Bacharach and Gambetta’s definition aptly depicts my friend’s initiation to big city life, this definition needs to be extended to capture additional forms of trust. Here, I consider two important deviations from their payoff matrix. First, Trusters might not always benefit from their actions, even if the Trustee is trustworthy. The pedestrian who accedes the stranger’s request to borrow his cell phone to make an urgent call, for example, is unlikely to receive any gain from his trusting behavior. This type of nonprofitable trust, in which short- or long-term material gains cannot motivate the Truster, has rarely been discussed or

---

\(^3\) Another influential definition of trust is offered by Coleman (1990: 99). He argues that trust constitutes a subclass of situations involving risk, namely, those interactions in which the risk entails the behavior of another actor. The significant difference between Coleman’s definition, which he presents algebraically, and the definition offered by Bacharach and Gambetta is that the preferences of the Trustee in Coleman’s definition are reduced to a probability. For Coleman, it might or might not be in the Trustee’s self-interest to be untrustworthy; for Bacharach and Gambetta, the risk for the Truster is that the Trustee can always gain by being untrustworthy.
examined, so it is not clear how frequently and under what conditions it occurs.\(^4\) In the above payoff matrix, it exists when the trusting action, A, is a dominated strategy, namely, \(x > 3, y > -3\). Rational choice and game theory, of course, predict that nonprofitable trust will not occur unless short-term potential losses can be recouped through future gains.

Second, encounters involving trust are often embedded in a series of interactions. Accordingly, it is worth distinguishing between repeated and nonrepeated trust. For many scholars, repeated interactions are essential for the emergence and stability of trust. For example, Dasgupta (1988: 49) maintains that “[f]or trust to be developed between individuals they must have repeated encounters, and they must have some memory of previous experiences.” Cook and Hardin (2001: 328) claim that the principal incentive to form trusting relationships is “the desire to maintain the relationship itself for future activities.” Intuitively, we assume that individuals will lend their cell phones more readily to a neighbor than to a stranger, since they are more likely to interact with the neighbor in the future.

Repeated trust interactions can take two forms. In some interactions, the roles of Truster and Trustee are reversed over time; in other interactions, the roles are maintained for the duration of the relationship or are rarely reversed. Coleman (1990: 177–178) refers to these types of trust as mutual and asymmetric trust, respectively. While the latter often characterizes the relationship between religious leaders (Trustee) and adherents (Truster), this paper will not explore this relationship, which deserves a separate comprehensive analysis, but rather will focus on the symmetric and asymmetric interactions between laypeople in religious communities.

Notice that trust has been defined without any reference to the costs and benefits of reputations or punishments. These factors, however, can be easily incorporated into the payoff structure and in fact may have already been assumed in the cell phone example above. Aside from the greater knowledge we have about the trustworthiness of our neighbor than about that of a stranger, we are also more inclined to loan a cell phone to a neighbor because we stand to gain reputational benefits from our trusting behavior and can impose greater punishments on a neighbor if he calls Thailand rather than his wife who is expecting him home for dinner. Incorporating reputations and punishments into

\(^4\) For instance, most economic experiments that are aimed at measuring trust assume, consistent with Bacharach and Gambetta’s definition, that the Truster can potentially benefit from a trusting relationship if the Trustee is trustworthy (e.g., Barclay 2004; Berg, Dickhaut, and McCabe 1995; Bohnet and Croson 2004; Croson and Buchan 1999; Glaeser et al. 2000; King-Casas et al. 2005; McCabe, Rigdon, and Smith 2003; McKelvey and Palfrey 1992). I refer to this type of trust as profitable.
the payoffs may actually “solve” the trust dilemma by explaining it away. For instance, reliable punishments can change the payoffs of an interaction so that trust is not a behavioral option. As Coleman (1990: 115) notes, “[i]f a trustee’s behavior is affected by the application of sanctions, then a trustor’s decision about whether to place trust in that trustee another time should be based not simply on his estimate of the probability of the trustee’s keeping the trust, but also in part on the use of negative sanctions.” If a Trustee can be punished for an untrustworthy decision (B’) such that he receives 0 instead of 4, the interaction no longer offers the possibility of trust, since it is in both the Trustee’s and Trustor’s self-interest to act cooperatively. Similarly, if Trustees can gain reputational benefits from trustworthiness, it may be in their self-interest to choose B instead of B’, thus eliminating the trust dilemma.

Now that we have our working definitions of trust, let me clarify the primary question of this article, which was offered in its title. To paraphrase, with admittedly less concision: Are individuals within a religious community more likely to make a decision that conforms to the payoffs defined by Bacharach and Gambetta as trust than individuals who are not in a religious community? When given the opportunity, the answer to this question is likely to be “yes.” But more commonly, especially in closed religious groups, these payoffs do not depict the range of prosocial interactions that religious adherents face within their communities. Closed religious groups typically achieve their collective goals by altering these payoffs rather than by relying on trust. Paradoxically, it is in religious groups where membership is fluid, demands are fewer, and the structure is less formal that trust as defined above is more commonly observed.

**RELIGIOUS SIGNALING, COOPERATION, AND TRUSTWORTHINESS**

Anthropologists have long noted that one of the primary functions of religion is the promotion of group solidarity, and most have recognized ritual as the...
mechanism through which this solidarity is achieved. Guided by Durkheim (1995 [1912]), who was among the first to appreciate the unifying nature of religious ritual, anthropologists have argued that collective rituals enable the expression and reaffirmation of shared beliefs, norms, and values and are thus essential for maintaining communal stability and group harmony (e.g., Douglas 1966; Radcliffe-Brown 1952). For Durkheim, collective rituals are the means by which individuals bond with one another in the community. He claims that the “effervescent” state of ritual performance minimizes individual distinctions and emphasizes the unity of the group. Turner (1969) also views this “effervescent” state as central to ritual’s efficacy. He observes that the temporary leveling of initiates during rites of passage increases a sense of communitas among ritual performers, which he characterizes as a strengthening of social bonds and heightened communion. As Stark (2001: 183) succinctly concludes, “Participation in religious rituals builds group solidarity.”

More recently, evolutionary scholars have argued that social bonding is not an end in itself but a means to facilitate intragroup cooperation (Bulbulia 2004b; Cronk 1994; Irons 2001, 2004; Nesse 1999; Rappaport 1999; Sosis 2003; Sosis and Alcorta 2003; Steadman and Palmer 1995; Wilson 2002). Large-scale cooperation, however, is notoriously difficult to attain without social mechanisms that limit the potential to free ride on the efforts of others (Dawes 1980; Hardin 1982; Olson 1965). The core challenge in achieving cooperative goals is solving problems of trust and commitment (Frank 1988, 2001; Schelling 1960, 2001). Intragroup cooperation is most likely to emerge when individual group members can guarantee their participation. The dilemma in human social interactions is how to generate a credible guarantee—in other words, one that can be trusted. Simply advertising a willingness to cooperate is fraught with potential deception. When faced with the conditions of collective action, the incentive to display false commitment signals is especially high, since individuals can achieve their greatest gains by refraining from cooperation while others cooperate. Therefore, whenever an individual can achieve net benefits from defection, credible signals of cooperative intentions tend to be those that are too costly for defectors to imitate (Spence 1974; Zahavi and Zahavi 1997).

Anthropologists (Cronk 1994; Irons 1996a, 2001; Sosis 2003), economists (Berman 2000; Carr and Landa 1983; Iannaccone 1992, 1994), and cognitive scientists (Atran 2002; Bulbulia 2004a, 2004b) have independently suggested that religious practices and emotions are indeed such signals. Observers of religion have repeatedly noted the costliness of religious obligations. Religious behaviors often entail significant proximate costs, such as time, energy, and material costs, as well as physical and psychological pain. For example, many populations require males and females to undergo initiation rites that include beatings, genital mutilation, exposure to extreme temperatures, tattooing, isolation, food and water
deprivation, consumption of toxic substances, and death threats (e.g., Tuzin 1982; Whiting, Kluckhohn, and Anthony 1958; Young 1965). In literate societies, religious legal codes (e.g., Laws of Manu, Shari’a, Talmud) outlining appropriate behavior tend to be formalized and regulate a wide range of activities, including food consumption, work, charitable obligations, and dress. These codes also define the frequency and structure of ritual ceremony and prayer. Although there may be physical or mental health benefits associated with some ritual practices (see Levin 1994; McCullough et al. 2000; Reynolds and Tanner 1995), the significant time, energy, and material costs involved in imitating such behavior serve as effective deterrents for anyone who does not accept the teachings of a particular religion. Therefore religions often maintain intragroup solidarity by requiring costly behavioral patterns of group members. The performance of these costly behaviors signals and engenders commitment and loyalty to the group and the beliefs of its members. Trust is enhanced among adherents, thereby facilitating cooperative pursuits.

Evolutionary models offer further insight into the efficacy and persistence of costly religious practices. Bliege Bird and Smith (2005) outline four necessary conditions for the evolutionary stability of a costly signal in a population: (1) There is within-group variance in some unobservable attribute, (2) group members can benefit from reliable information about this variance, (3) signalers can achieve benefits at the expense of those receiving the signal, and (4) the cost or benefit to the signaler of sending the signal is correlated with the signaler’s quality of the attribute. Sosis (2003) has argued that religious behaviors meet these conditions: (1) The intensity of religious beliefs varies within communities, and this variance is unobservable; (2) individuals benefit from accurate information about this variance because intensity of belief is related to one’s commitment to the group and its goals; committed members are more likely to be cooperators and thus preferred social interactants; (3) religious groups offer various benefits for in-group members that are mutually provided and are at risk of exploitation by those not committed to group goals; and (4) the cost or benefit of ritual performance is weighed against ritual opportunity costs that are expected to be higher for skeptics than for believers. Thus religious behavior can be understood as a costly signal that reliably advertises the unobservable condition of religious belief and group commitment. The time, energy, material, and opportunity costs of religious activity serve to prevent those who lack sufficient belief from displaying the signal.

While prior comparative, ethnohistorical, and experimental studies (Ruffle and Sosis n.d.; Sosis 2000; Sosis and Bressler 2003; Sosis, Kress, and Boster 2004; Sosis and Ruffle 2003, 2004) have been largely supportive of the application of costly signaling theory to religion, this research has focused exclusively on demonstrating a relationship between collective ritual practice and
cooperative behavior. None of this work has assessed whether trust mediates the observed relationship between costly religious practice and cooperation. Nor has the pioneering club goods model of religious behavior developed by economists, and its concomitant evaluation, focused on this relationship (e.g., Berman 2000; Iannaccone 1992, 1994). To address this lacuna, below I examine the role of trust in fostering and sustaining cooperation in several well-studied religious populations.

As was noted above, many scholars assume that religion promotes trust, thereby facilitating collective action, yet the lack of specificity about the meaning of trust has resulted in numerous claims that may be unwarranted. Debates about the existence of trust in bucolic areas (Amato 1993) are instructive. Consistent with the idyllic stereotype, some researchers (Fukuyama 1995; Putnam 2000) attribute high levels of trust to small towns in rural America. Cook and Hardin (2001), however, argue that small communities achieve high levels of cooperation and reciprocity through repeated interactions, direct information sharing about individual behavior, monitoring, and sanctioning. In a small community, the threat of ostracism is significant, especially if the costs of leaving the community are high (e.g., substantial land or business investments, or extensive family or social ties in the community) or there are barriers to entering other communities. If migration is costly or the threat of punishment looms, prosocial reciprocal interactions can be rational and strategic, not requiring any trust.

Rappaport (1999: 204) offers similar observations about various indigenous populations:

In tribal societies, ethics are an immediate and perceptible aspect of relations among people who are, for the most part, not only known to each other but stand in well-defined relationships to each other. Reciprocal (although not necessarily symmetrical) obligation is the cement if not, in fact, the ground of all such relationships, and the obligations they entail are usually quite clearly specified. Violations of obligation inevitably become evident, often quickly, and sanctions against breach of obligation are essential elements of reciprocity’s fundamental structure. … That prestige as much as or even more than wealth is among the chief rewards of life properly lived in societies in which reciprocity prevails also encourages vigorous, valorous and generous fulfillment of obligation.

---

7 Consistent with these results, Orbell et al. (1992) found that church attendance among Mormons in Logan, Utah, where over 75% of the population are members of the Church of Latter-day Saints, was positively correlated with cooperation toward anonymous strangers in prisoner’s dilemma experiments. In a more religiously diverse area, no correlation was found, suggesting that ritually generated cooperation is aimed at fostering intragroup prosociality rather than generalized or universal prosociality.
Communities that are defined by their religious identity may often achieve cooperation through similar means. Although contemporary world religions have members dispersed across the globe, adherents generally live in small, well-defined groups that are distinguished by their collective performance of rituals (Luhmann 1997; Wilson 2002). Many religious communities, past and present, have established social control through the threat of ostracism as well as formal punishment. As in tribal and rural communities, adherents develop long-term relationships and are able to observe the behavior of other residents, thus possibly solving the trust dilemma through reputational rewards and punishment threats that alter the payoffs of social interactions so that trust is unnecessary.

It is ostensibly in the economic sphere that religious groups have benefited most from the high levels of trust attributed to them. There are countless examples of small religious groups that have profitably dominated distinctive economic niches that require considerable trust and cooperation (e.g., Landa 1994). I turn to three such populations—Jewish diamond merchants, Maghribi traders, and Muslim traders in Africa—to assess how they overcome the considerable trust dilemmas posed by their commercial activities. Each population has been widely studied and offers an opportunity to examine the relationship between signaling, trust, and cooperation.

**Jewish Diamond Merchants**

The business dealings of Jews in diamond centers such as Antwerp, Amsterdam, London, New York, and Tel Aviv are often cited as a classic example of trust (e.g., Coleman 1990; Gutwirth 1968; Putnam 2000; Starr 1984). For instance, Wechsberg (1966) describes buyers in London walking around “with a hundred thousand pounds worth of diamonds that were handed over to them in trust” (cited in Coleman 1990: 109). Despite the opportunity for theft and fraud between buyers and sellers, “[a]ll deals are settled verbally” (Coleman 1990: 109). In a rare ethnographic study of New York’s diamond district, 8 Shield (2002: 103–105) describes the extraordinary level of trust among Ultra-Orthodox Jews; “handshakes, Yiddish, and trust still close multi-million dollar deals” (Shield 2002: 1).

As was defined above, trust therefore appears to be profitable, repeated, and either mutual or asymmetric, depending on the specific relationship. However, Williamson (1993: 471) argues that the “appearance of trust among diamond dealers is deceptive.” Williamson follows Bernstein (1992), who claims that trust is not critical or even important for business activity among New York’s Jewish diamond dealers. Instead, reputation and social bonds are backed up by a private

---

8 New York’s diamond district is “more than 95% Jewish” (Shield 2002: 15).
arbitration system that uses reputation and the threat of expulsion to maintain low-cost commercial transactions. Williamson asserts that Jews succeeded in the diamond business not because of trust, but because they were “able to provide cost-effective sanctions more efficiently than rivals” (1993: 472).

In a recent analysis, Richman (2004) questions why diamond merchants have rejected public courts in favor of a private internal adjudication system. He argues that the unique nature of the diamond industry, namely, regularly handling small, extremely valuable goods that can be easily resold in distant markets, makes the public legal system incapable of enforcing contracts between diamond sellers and buyers. While Bernstein (1992) notes the importance of reputations supported by the internal adjudication systems, Richman recognizes that the high value of diamonds makes one-time theft more profitable than long-term benefits accrued from an honest reputation, thus making a scrupulous business reputation alone insufficient to maintain reliable contracts and exchanges. It is more likely that the added social costs of defection within tight-knit Ultra-Orthodox communities enables these communities to maintain their competitive edge in this industry by keeping transaction costs low. In other words, because ruined reputations affect not only business opportunities but also one’s social life, including diminishing the marriage prospects of one’s children and siblings, these threats help to maintain cooperative exchanges. Whether meted out socially or through the private court system of diamond traders, these costs alter the payoffs of transactions, so it is in everyone’s interest to pursue honest business dealings. Therefore, the success of Jewish diamond merchants does not appear to depend on any form of trust.

Maghribi Traders

Many scholars have claimed that medieval trade flourished as a consequence of trust relations successfully established among “natural groups,” including religious communities (e.g., Landa 1994; Rosenberg and Birdzell 1986; Sombart 1953; Weber 1961). Maghribi traders, a group of 11th to 12th century Jewish merchants who conducted commerce extensively throughout the Mediterranean, would seem to be a case in point (Greif 1989). All medieval merchants faced a significant quandary: How should they manage their goods abroad? If they traveled with their merchandise, this would limit the amount of business they could conduct and limit their ability to diversify their trading activity, a useful strategy to buffer against failed ventures. However, if they hired an agent to manage their goods abroad, they could not observe the transactions and would have to rely on the agent to honestly report the selling prices of the goods. Relying on the courts to punish dishonest agents was an ineffective deterrent, since the costs to monitor such transactions were too high. To overcome these
difficulties, Maghribi traders arranged agency relations through a peer organization, which Greif (1989, 1993) refers to as a coalition. Greif argues that the coalition was successful at maintaining honest and trustworthy agents because the short-term gains that could be achieved from cheating (e.g., stealing goods, inaccurately reporting a sale price) were outweighed by the long-term gains achieved by remaining in the coalition.

Greif (1989) refers to the merchant-agent relations as one of trust, but it is evident that the payoffs defined above do not characterize these relations. Agents who cheated a merchant were punished by a permanent refusal to conduct future business with all members of the coalition. Hence upholding an honest reputation was crucial for remaining in the coalition and receiving its benefits. Both merchant and agent were aware of the payoffs involved: “The agent required a reputation of being honest, and the merchant could trust him” (Greif 1991: 460). Yet given the institutional framework, reliance on reputations, and effective means of punishment, Maghribi merchants did not need to trust their agents. As Greif argues, the agents’ payoffs for honesty outweighed their payoffs for defection. Greif (1989: 881) further contends, “trust did not reflect a social control system or an internalization of norms of behavior.” He rejects the possibility that trust between Maghribi traders was maintained through reputations signaling trustworthiness “because he fears God, or has internalized an ideology of honesty” (Greif 1989: 867). Rather, Greif claims, the successful merchant-agent transactions are explained by the reputational mechanism and credible threats of punishment.

**Muslim Traders in Africa**

Muslim merchants have been credited with the spread of Islam throughout Africa and Southeast Asia, including Indonesia, the Philippines, and Malaysia (Stark 2001:84). Focusing on Africa, Ensminger (1997) convincingly argues that conversions resulted from the economic advantages that Islam offered to traders (see also Triulzi 1975). She claims that “Islam was a powerful ideology with built-in sanctions which contributed to considerable self-enforcement of contracts. True believers had a non-material interest in holding to the terms of contracts even if the opportunity presented itself to shirk” (Ensminger 1997: 8). Thus Islam provided a mechanism to overcome the collective action problems of long-distance commerce. According to Ensminger, conversion to Islam increased trust among traders, which reduced transaction costs, making trade more profitable. In addition, high levels of trust among Muslim coreligionists allowed for greater credit to be extended, thereby facilitating further trade expansion. Similar to the

---

9 According to the definitions offered above, this trust would be profitable, repeated, and asymmetric.
case for Jewish diamond merchants and Maghribi traders, trust for Muslim traders is characterized as profitable, repeated, and either mutual or asymmetric. Ensminger contends that the steep initiation costs of entry into Islam, such as daily prayer, abstaining from alcohol, fasting during Ramadan, and the pilgrimage to Mecca, served as the means for establishing a reputation among merchants for trustworthiness. In other words, these rituals and taboos are costly signals of commitment that served to prevent free riders from achieving the benefits of more efficient commerce.

However, Muslim traders did not rely exclusively on “non-material interests” to secure cooperative gains. As Ensminger (1997: 9) notes, “one of the most important institutions spread through the trading diasporas was the Muslim commenda.” The commenda was nearly identical to the coalition used by Maghribi traders discussed above. Perinbam (1980, cited in Ensminger 1997: 10) describes the commenda:

> It enabled an investor or group of investors to entrust capital or merchandise to an agent-manager, who traded with it and returned the principal and previously agreed-upon share of profits to the investor. The agent received the remaining share of profits as a reward for his time and labor. … The commenda, which combined the advantages of loan with those of partnership, was a contract which regulated the use of capital, trading skills, and labor for mutual profit.

As with Maghribi traders, trust was not critical in agent-trader interactions because agents could maximize their gains through honesty. The necessity of an honest reputation to remain in the commenda altered the payoffs of the trust dilemma such that it was not in the agents’ self-interest to defect on the arrangement.

**Why Trade with Coreligionists?**

Many trade networks, all with their specific collective action problems to solve, are dominated by religious communities (e.g., Iyer 1999; Udovitch 1970). However, if Jewish diamond merchants and Maghribi and Muslim traders exemplify conditions of other religious trading communities (i.e., they solve the trust dilemma by modifying the payoff structure), why have these business networks been organized around religious identities? Indeed, if the payoffs are structured so that reputations are valued and honesty is in everyone’s interest, why does it matter whether your trading partner shares your religious beliefs and practices?

Three factors appear to be important. First, social relations within religious communities are circumscribed. Ritual obligations and symbolic markers demarcate membership within religious communities, thus clearly defining the
social arena and with whom social obligations lie. By establishing an unambiguous social environment, religious groups are able to offer a history of social interactions on which reputations can be built and evaluated. In other words, the religious identities of Jewish diamond merchants, Maghribi traders, and Muslim traders do not directly signal honesty; what keeping *kashrut* or *hallal* (Jewish and Muslim dietary laws respectively), pilgrimages, praying three or five times daily, wearing a yarmulke and other religious obligations signal is the ability to obtain reliable reputational information because past social relations will be common knowledge within these respective communities. It is easier to track individuals and obtain information about their past interactions when community membership is well-defined; religious behavior and symbolic markers visibly delineate community boundaries. The second reason why successful collective action so often involves members who share a religious identity is that punishment can be extremely effective when it not only affects economic associations, but affects social relations as well. Similar to the Jewish diamond merchants who suffer social costs if they cheat on a transaction, Maghribi traders would impose social sanctions (in addition to trade sanctions) on defecting agents by appealing to the entire Jewish community (Greif 1989). Third, supernatural sanctions (e.g., hell) and rewards (e.g., heaven) can alter the perceived payoffs of an interaction so that the trust dilemma is avoided (see Bulbulia 2004b; Johnson and Kruger 2004; Sosis 2003). Interestingly, supernatural sanctions and rewards are not relied on exclusively to promote cooperation; they are used to support and strengthen material punishment and reputational mechanisms. To clarify, I am not positing that either of these systems necessarily has priority over the other. The point is that one reason religious groups successfully overcome the challenges of collective action (e.g., Sosis and Bressler 2003) is because their material and supernatural punishment and reward mechanisms reinforce and complement each other.

**DOES TRUST EVER EXIST?**

While the examples discussed above have been limited to economic activities, many religious communities employ punishment and reputational mechanisms to circumvent trust dilemmas by altering the payoffs for a wide range of within-group social interactions. The 19th century Perfectionists of John Humphrey Noyes’s Oneida community, for example, formalized the reputational mechanism through the practice of mutual criticism, in which members stood in front of the community and endured humiliating public critiques of their character, typically highlighting spiritual and social shortcomings (Oved 1988). Of course, many religious groups, such as the Amish and the Church of Latter-day Saints, hold the threat of excommunication over their constituents, a threat that they can reliably
enforce (Hostetler 1980; Shepherd and Shepherd 1998). Social sanctions, however, need not be formalized. Although Jehovah’s Witnesses will excommunicate those who blatantly deviate from the Society’s teachings, a process referred to as disfellowship, informally the costs of not fully accepting the belief system is a “loss of interaction and the cutting off of affective bonds” (Zellner 2001:358), which effectively results in the social excommunication of the nonbeliever.

In closed religious communities, evidently, recurrent cooperative pursuits often do not meet the conditions for trust defined by Bacharach and Gambetta (2001). When social relations occur within the framework of a bounded community so that interactions are repeated and mobility between social groups is limited, reciprocity can emerge and stabilize. Consequently, reputations and punishment threats become relevant and influence the payoff structure of the interactions, often making trust unnecessary. Nonetheless, trust does indeed exist within religious populations, and I now turn to the conditions under which religious signaling is an important factor in promoting trust.

**Monitoring Costs**

Although repeated interactions, reputational effects, and the threat of ostracism all facilitate cooperative interactions, religious rituals and markers may promote trust within religious communities by reducing monitoring costs of collective pursuits. For example, in fieldwork that I conducted among Ultra-Orthodox Jews in northern Israel, I found monitoring costs of cooperative activities to be especially low. I regularly attended community religious functions at which bags and valuables lay unguarded in a corner. Following these events, I occasionally visited the local university library, where, in sharp contrast to the gathering I had just left, I was required to lock my belongings in a locker, and observation cameras monitored the book stacks and exits.

Religious groups may achieve low monitoring costs in collective pursuits, but they pay additional costs to monitor the performance of religious obligations. However, since many religious behaviors are formal, conspicuous, and communally practiced—indeed, these characteristics contribute to their effectiveness as signals (Alcorta and Sosis 2005; Sosis 2003)—they are cheaper to monitor than other activities. Thus, religious groups lower their net monitoring costs by transferring the costs from monitoring collective inputs to monitoring conspicuous religious activities.

Monitoring costs are directly reduced by three additional means, all of which increase religious group members’ opportunity costs, making defection and untrustworthy decisions less likely. First, the high investment in adolescent religious training rather than secular education, such as among Hutterites
(Hostetler 1997) and Ultra-Orthodox Jews (Heilman 1992), may make entry into other communities difficult if individuals lack skills to compete economically (Sosis 2003). Second, permanent stigmatizing markers such as tattoos, piercings, or more severe bodily mutilations may limit one’s ability to join an outside group. Third, Sosis (2003) has argued that through processes of cognitive dissonance (Festinger 1957) or social perception (Bem 1965, 1972), repeated ritual performance fosters belief in the doctrines of the group; it facilitates internalizing group norms. These beliefs typically include a nonmaterial system of reward and punishment that consequently reduces the costs of monitoring members’ behavior (Bulbulia 2004b; Johnson 2005; Johnson and Kruger 2004). Alcorta and Sosis (2005) further note that humans use ritual to conditionally associate emotions and culturally defined sacred symbols; they propose that such symbols represent powerful tools for promoting trust and cooperation. Since emotions are generated from limbic structures that are out of conscious control, they are difficult to fake (Ekman, Levenson, and Friesen 1983, Levenson 2003) and can consequently serve as reliable signals of trustworthiness and commitment, thereby reducing monitoring costs (Alcorta and Sosis 2005; Bulbulia 2004b).

While not directly concerned with supernatural reward and punishment, Coleman (1990) makes two germane points for understanding the conditions under which a society will aim to internalize its norms. First, Coleman (1990: 296) notes that norm internalization is efficient when there is a range of actions that the society seeks to control. This aptly characterizes religious communities, which generally seek members who behave prosocially toward coreligionists under diverse conditions; in other words, religious communities wish to encourage cooperation and trust between members regardless of the situation that arises. Second, Coleman argues that external policing becomes less efficient when observations are more costly to observe, and therefore under these conditions, societies are more likely to rely on internalization strategies. Since the intragroup trust and cooperation that are encouraged within religious communities are not limited in time (e.g., just during work hours) and place (e.g., just in a house of worship) but are a continuous obligation, it is impossible to monitor members’ commitment to this ethic all of the time. Johnson and Bering (n.d.) additionally note that negative sanctions are more effective than positive reinforcement in eliciting a desired behavioral response. Indeed, reviewing the results of various economic experiments that employed reward and punishment incentives, Andreoni, Harbaugh, and Vesterlund (2003: 894) conclude that “one might expect less cooperation in societies where good behavior is rewarded than in those where poor behavior is punished.” Hence, supernatural punishment is an efficient complement to the material punishment system that deters defection from cooperative obligations among community members, resulting in lower monitoring costs.
Returning to Bacharach and Gambetta’s (2001) definition, low investment in secular training, stigmatization, and norm internalization make it more likely that a Truster will choose $A$, the trusting decision, since she recognizes that Trustees are more likely to choose $B$ because the potential gains for defection ($B'$) are reduced, or at least perceived to be reduced, in the case of supernatural sanctions.

**Intergroup Mobility**

Exponential population growth, greater demographic mobility, and the trend toward increased urbanization in the modern era have influenced the scope of trust in contemporary societies. The effects of these factors on trust might seem counterintuitive; where social groups are open and fluid, there will be greater dependence on trust in collective pursuits than there is in segregated or isolated closed communities. In environments in which there is high mobility between groups, reputations and punishments are less effective as means of securing cooperation (Hardin 2000). Since individuals can relocate and align themselves with a different community when faced with the prospect of social sanctions, punishments are less effective in generating cooperation. Likewise, mobility enables individuals to escape and leave behind damaged reputations, making reputations in open groups a less reliable source of information than in closed communities. Ironically, trusting behavior may be most vital in environments in which social groups are fluid and individuals maintain multiple group identities. Under these conditions, cooperation is less frequent and more challenging to sustain than in closed groups (e.g., Sosis and Bressler 2003); however, when collective goals are pursued, trusting behavior is essential. This is not to claim that in open communities with high intergroup mobility, reputations are not valued and punishments are not employed with some efficacy. They are. The point is that these mechanisms for motivating collective action will be less effective in open communities than in closed ones. Concomitantly, in open communities, the payoffs for cooperative interactions are more likely to mirror the conditions for trust established above; hence cooperation is more difficult to achieve. In other words, open communities rely on trust to achieve their collective goals to a greater degree than do closed communities, which in contrast are more dependent on institutional structures that enable punishment and reputational mechanisms to foster cooperation.  

Religion plays another significant role in environments with high intergroup mobility. Individuals who maintain multiple group identities can face conflicts of interest when diverse group ideals do not coincide. In these circumstances,

---

10 To clarify, I am not arguing that individuals’ general perceptions or attitudes of trust will be higher in open communities than in closed communities. Indeed, Alesina and La Ferrara (2002) find that interpersonal trust is somewhat lower in communities where members are more transient.
religion often serves to prioritize values. Theologians (Hefner 1993; Tillich 1951, 1952) as well as social scientists (Klass 1995; Rappaport 1999) have noted that one of religion’s defining characteristics is that it is the ultimate concern of its followers. As such, religions not only offer moral guidelines to settle individual conflicts of interest, but also establish a hierarchy of values that can be incorporated into other identities. This integration will be facilitated by the fact that others equally recognize the priority of religious values and convictions.

**Nonrepeated Trust**

Costly religious rituals and markers not only lower monitoring costs of community members and encourage repeated trust in environments of high mobility; they also serve to enhance trust among fellow adherents in the absence of long-term reciprocal relations. These interactions were characterized above as nonrepeated trust and include the payoffs that I defined as nonproftitable, as well as profitable trust. In large-scale societies, generalized norms of cooperation are susceptible to free riders because the ease of relocating to another community makes punishment and reputational benefits ineffective means of social control (Hardin 2000). In addition, higher population densities result in more frequent interactions between individuals who do not have long-term reciprocal relations. Costly rituals and stigmatizing markers enable mutually beneficial relationships between strangers or new acquaintances without the costly investment of long-term relationships. These religious obligations are likely to be most valuable in communities with adherents in separate locations (whether geographically distant or near) who interact occasionally but not frequently enough to establish cooperative networks without the support of additional commitment signals. Trade, marital exchanges, and financial aid to support struggling and nascent communities are all realized benefits of these intercommunity relations between dispersed communities that share a common faith (e.g., Ensminger 1997; Hostetler 1997; Udovitch and Valensi 1984).

While religion undeniably facilitates collective action among community members, its greater influence may be the proliferation of nonrepeated trust between members of the same religion who are strangers (see Macy and Skvoretz 1998), and it likely offers the uncommon conditions under which nonproftitable trust can occur. In contrast to intragroup or intergroup trust, we can refer to this type of trust as extended-group trust: trust that occurs among anonymous (or nearly anonymous) members of the same religion who reside in different communities. Computer simulations conducted by Riolo, Cohen, and Axelrod (2001) show that cooperation can emerge and stabilize between agents that share identifying markers and have no prior history of interaction and no expectations of future encounters. Costly religious markers and rituals enable adherents who
share a common faith to assess mutual trustworthiness without investing in long-term reciprocal relations. For example, Triulzi (1975: 59) describes how 19th century migrant Muslim traders from the Sudan traveled to their coreligionists in Ethiopia, where “[t]he stranger was received as a brother.” Similarly, during fieldwork among Ultra-Orthodox Jewish communities in Israel, I repeatedly observed invitations for meals, lodging, and rides by residents to unknown Ultra-Orthodox travelers. On several occasions, I witnessed cars being loaned to complete strangers, and interviews revealed a surprising number of interest-free loans offered and accepted between people who had previously not known each other.

It is puzzling why such markers are perceived as honest signals of trustworthiness. To imitate nonpermanent markers with intentions of deception is generally not too difficult. Palestinian suicide bombers, for instance, have often dressed up as Ultra-Orthodox Jews, escaping detection and successfully fulfilling their missions. With less devastating effects, reporters, anthropologists (including myself), and Peace Corps workers, among others, have regularly donned the garb of those they are working with to gain their trust. Symbolic markers have long been recognized as critical for establishing ethnic and cultural boundaries, allowing individuals to selectively direct altruistic benefits toward in-group members (Nettle and Dunbar 1997). Barth (1969) accurately notes, “The identification of another person as a fellow member of an ethnic group implies a sharing of criteria for evaluation and judgment” (cited in Carr and Landa 1983:150). We appear to have psychological biases that favorably predispose us toward others with whom we share symbolic identity markers (e.g., Fershtman and Gneezy 2001). In our evolutionary history, such markers were likely to be reliable indicators of in-group identity, as they generally, but not always, are today. Despite these possible biases, when reputational information is available, it is rarely ignored; it is evaluated along with symbolic messages, often taking priority. For instance, in her study of Ultra-Orthodox diamond merchants, Shield (2002: 110) discerns that “religious piety does not substitute for ethical business practice; in fact, expressions of piety without ethical behavior garner considerable contempt. Traders are too savvy to let superficial emblems of religiosity blind them to dubious marketplace behaviors.” Jains from India offer another interesting example. Iyer (1999: 107) observes that among Jains, “business reputation is derived not only from economic activity but also from religious activities of family members.” However, religious practices appear to play a signaling role that is secondary to economic status. In Jainism, wealth is considered a reward for one’s piety, whereas involuntary poverty is a result of moral failings. Consequently, “wealth emerges as a signifier not only of moral piety but also as the signal for economic reputation” (Iyer 1999: 105). Wealth begets the advantages of extended credit, which of course begets more wealth.
Signaling Trustworthiness to Non-Group Members

Few studies have examined how signals that may have emerged for one purpose can be coopted for other purposes and still maintain stability (e.g., Johnstone and Norris 1993). With regard to religious symbolic markers, it appears that they not only signal in-group commitment and trustworthiness; outsiders also use religious behaviors and markers to evaluate trustworthiness. For example, Frank (1988) observes that affluent New York City families place advertisements in the newspapers of Salt Lake City for Mormon governesses for their children. Apparently, “persons raised in the Mormon tradition are trustworthy to a degree that the average New Yorker is not” (Frank 1988: 111).

Paxson (2004) has similarly shown that Sikhs achieve economic benefits from perceived trustworthiness. Adult male Sikhs can be easily distinguished by various external markers, often referred to as the five K’s (Kes, Kangha, Kara, Kirpan, and Kachera): unshorn hair and beard and wearing a comb, steel bracelet, saber, and breeches. Additional constraints on their behavior, such as refraining from alcohol and tobacco and the requirement to pray five times daily, further mark their distinctiveness. Paxson (2004) argues that Sikhs are recognized by non-Sikhs as trustworthy trading partners, even without a history of prior exchanges. Non-Sikhs can utilize Sikh symbolic markers as a “seal of approval” signaling trustworthiness. The external displays indicate that the individual has already endured the monitoring systems within Sikh communities that allow him to maintain his membership.

CONCLUSIONS

Religious communities clearly achieve high levels of cooperation and collective action. Trust, however, is often unnecessary to mediate these cooperative activities. Closed religious communities tend to support effective punishment systems, be well defined, and have members who repeatedly interact with each other, enabling the establishment of reliable reputations. These religious communities additionally facilitate collective action by offering a credible history of social relations and extending their punishment measures outside of the economic sphere and into the social arena. In religious communities where identities are shared between multiple groups and individuals can transfer groups, reputations and punishments will be less successful at marshaling cooperation, and trust will be indispensable for realizing cooperative goals. Religious markers and rituals also promote successful trust between anonymous interactants. Not only do religions achieve this among extended-group members, that is, those who
share a faith but live in disparate communities, but religious rituals and markers can also signal trustworthiness to non-group members. Nonetheless, religious adherents are unlikely to extend trust indiscriminantly. Indeed, various researchers argue that in social systems with close-knit families or small groups, the strong trusting relations that occur within the family or group are at the expense of trust with individuals outside of these relations (Fukuyama 1995; Gellner 1988; Ruffle and Sosis 2005; Yamagishi 1996). Experimental evidence appears to support this assertion (Yamagishi and Yamagishi 1994; Yamagishi, Cook, and Watabe 1998). Among religious communities that engage in high levels of costly signaling and encourage strong group ties (even in high-population-density environments), we would similarly expect members to be less trusting of outsiders than are individuals living in large urban settings characterized by networks of reciprocation (see Cook and Hardin 2001).

There may be conditions, however, under which indiscriminant, or less discriminating, trust is a successful strategy. For example, following Weber, Runciman (2005) argues that the success of the Puritan ethic was that indiscriminant fairness and cooperation were profitable because in small communities, there were significant public pressures and enforceable punishments that maintained this norm of behavior. Simulations conducted by Macy and Skvoretz (1988) indicate that trust can spread across a large population if trust is first stabilized in embedded social interactions.

The discussion presented here raises a number of questions and suggests several avenues for further inquiry. Within the anthropological signaling literature (e.g., Irons 2001; Sosis 2003), it has been assumed that religious behaviors directly signal group commitments. However, the case studies examined here suggest that religious signals have multiple effects and that their messages and significata vary environmentally in fundamental ways. For example, in addition to signaling group commitment, within closed groups, religious signals can also indicate membership in a community where reputational information is valued and available. The prevalence of punishments and reputations in many of these communities to which high levels of trust have often been ascribed attests to the importance of these mechanisms in facilitating cooperative goals. When given the opportunity to behave in a trusting manner according to the payoffs defined above, individuals within these communities will undoubtedly often rise to the challenge. What is significant is that many religious groups do not rely on the goodwill that their communities cultivate but rather establish institutions that strengthen the efficacy of social penalties and reputational status. As Sigmund, Hauert, and Nowak (2001: 10757) conclude from their analysis of minigame simulations of public goods experiments, “reputation is essential for fostering social behavior among selfish agents, and … it is considerably more effective with punishment than with reward.” Indeed, recent experimental work has found
reputations (King-Casas et al. 2005) and punishment (Fehr and Gachter 2002) to be powerful predictors of trust and cooperation. Future work should further explore the efficacy of these mechanisms within religious communities and the relative costs of their maintenance.

Future research also needs to examine trust relations in open religious communities. I have asserted here that collective action in these communities will require trust owing to weak punishment and reputational mechanisms, yet empirical data are lacking. Observational as well as experimental data are imperative for evaluating this hypothesis. Moreover, future work must examine trust within religious populations that constitute the majority in their respective environments. When minority coreligionists from disparate communities interact, we expect markers and ritual signals to facilitate social interactions. Are such signals as effective when coreligionists constitute the majority of the religious landscape? In the United States, for example, is wearing a Star of David more effective at fostering trust among Jews than wearing a cross is at fostering trust among Christians? While experimental data could evaluate this, it seems likely that when coreligionists constitute a majority, they will require additional identity information to narrow the individual’s affiliation to a specific subpopulation within the majority culture. In contrast, minority religious populations are probably less concerned with denominational affiliations when evaluating trust relationships with unfamiliar extended-group members.

I opened this article by noting that current world affairs highlight the importance of understanding the relationship between religion and trust. I close with the same message. As more destructive means of warfare continue to develop and the dangers posed by international conflicts increase, understanding the factors that promote and sustain trust may be more relevant now than at any time in history. Examining religion should be central to this endeavor; even when conflicts are not directly motivated by religious concerns, religion is often used as a catalyst to unify groups for military purposes (MacNeill 2004; Sosis 2004; Sosis, Kress, and Boster 2004). In addition to its relevance in developing conflict resolution strategies, religion’s ability to enhance trust may also be invaluable for promoting economic growth (Barro and McCleary 2003) and economic exchange in underdeveloped economies in which property rights and other law enforcement, judicial, and economic institutions are weak (see Ruffle and Sosis n.d.). Over a decade ago, Coleman (1990: 180) suggested “that combined empirical and theoretical work on the dynamics of trust in social relationships should be productive.” I of course concur and hope that as this work continues to develop, future research will also explore the dynamics of trust within religious communities.
REFERENCES


We find no association with trust but we do find a statistically significant relation with reciprocity. Individuals more active in religious organizations tend to reciprocate more than individuals who participate less, even though their trustiness on others is about the same as that of less religiously active people. Does Religion Promote Trust? the Role of Signaling, Reputation and Punishment.