Forthcoming in *Biology and Philosophy*.

Please do not quote without permission until published.

The Two Faces of Revenge

Moral Responsibility and the Culture of Honor

Abstract:

Retributive emotions and behavior are thought to be adaptive for their role in improving social coordination. However, since retaliation is generally not in the short-term interests of the individual, rational self-interest erodes the motivational link between retributive emotions and the accompanying adaptive behavior. I argue that two different sets of norms have emerged to reinforce this link: (1) norms about honor and (2) norms about moral responsibility and desert. I observe that the primary difference between these types of retribution motivators lies in where the normative focus is placed after an offense. In the first form of retribution, the normative focus is on the offended party. In the second, it is on the offender. Next, I show how each class of norms is well tailored to the particular features of the environment in which these forms of retributive behavior emerge. Finally, I consider some philosophical implications of these observations. I suggest that my account, if correct, would pose tough challenges for contemporary philosophical theories of moral responsibility and punishment.

1. Introduction: The Two Faces of Revenge[1]

[1] I am grateful to John Doris, Joshua Knobe, Don Loeb, Shaun Nichols, Manuel Vargas, and an anonymous referee at *Biology and Philosophy* for their comments on earlier drafts of this manuscript.

The following passages illustrate two different ways in which the retributive instinct can manifest itself. In the first Busquet (1920: 357-358) describes the disgrace that befell those on the island of Corsica who did not avenge a wrong in timely fashion:

> Whoever hesitates to revenge himself, said Gregorovius in 1854, is the target of the whisperings of his relatives and the insults of strangers, who reproach him publicly for his cowardice. In Corsica, the man who has not avenged his father, an assassinated relative or a deceived daughter can *no longer appear in public*. Nobody speaks to him; he has to remain silent.

The second is from Kant (1996: 158):

> But whoever has committed murder, must die. There is, in this case, no juridical substitute or surrogate, that can be given or taken for the satisfaction of justice…Even if a civil society resolved to dissolve itself with the consent of all its members—as might be supposed in the case of a people inhabiting an island resolving to separate and scatter themselves throughout the whole world—the last murderer lying in prison ought to be executed before the resolution was carried out. This ought to be done in order that every one may realize the desert of his deeds, and that blood-guiltiness may not remain upon the people; for otherwise they might all be regarded as participators in the murder as a public violation of justice.

This paper explores the differences between these two attitudes regarding retribution, and what these differences mean for the contemporary philosophical debate about moral responsibility and punishment. After outlining what I take to be the philosophical significance of this analysis, I make some general remarks about the adaptive benefits of retributive emotions. I then examine the normative systems that reinforce the two forms of retribution described above—norms about

honor in the first case, norms about moral responsibility and desert in the second—and show how each system is well tailored to the particular features of the environment in which these varieties of retribution emerge. Finally, I provide a more detailed (albeit still too brief) examination of the social and philosophical implications of these observations.

2. Philosophical Significance

Philosophical theories of moral responsibility and punishment are numerous and diverse, but virtually all of them share two features in common. First, they present conditions that must be met in order for blame or punishment to be justified, conditions that are designed to apply universally. By this I don't mean that the theories conclude that all (or even any) adults in a given society can be appropriately blamed or punished. It may be that some or all members of a particular culture fail to meet the theory's criteria for justified blame or punishment. But the conditions or criteria *themselves* are meant to apply across cultures in such a way that *if* the sufficient conditions specified in the theory are met, the blame or punishment is justified; and if a necessary condition is not met, the blame or punishment is unjustified (unless there are other overriding moral considerations).[2] Second, the theories appeal to the reader's intuition in order to justify the criteria or conditions employed by the particular theory.[3] These two features make

[2] See Knobe and Doris (forthcoming) for an excellent examination of the ways in which theories of moral responsibility are structured.

[3] Of course, these features are not exclusive to theories of moral responsibility. Nichols et al (2003) show how premises in arguments for epistemological theories appeal to intuitions which may differ across cultures. Machery et al (2004) make a similar point about theories of references. And Stich (1988) argues that the problem of cognitive diversity threatens the epistemological tradition in general. My tentative conclusions in this paper are in many ways analogous to those in Nichols et al (2003). Indeed,

theories of moral responsibility vulnerable to an empirical challenge. If I am right (1) that honor cultures and non-honor cultures have radically different perspectives regarding responsibility and punishment, and (2) that these differences cannot in all cases be traced to irrationality on either side, then it is unlikely that a theory will be able to capture the considered intuitions of individuals in both cultures. It would therefore be more reasonable to regard our beliefs attitudes as artifacts of particular cultural and evolutionary histories, rather than as guides to the *truth* about the proper conditions for blame and punishment. I return to this discussion in the final section of the paper, after making the case that there are indeed radically different perspectives on retribution, blame, and punishment across cultures.

3. Why are we Retributive?

Retributive behavior is in a certain sense irrational.[4] Retaliation cannot undo the harm committed by the offense and it often comes with a significant cost or risk. Yet every known culture features retribution in some form, as well as norms and beliefs that govern and justify revenge behavior. One common explanation for this is that retributive dispositions are thought to be fitness enhancing. Consequently, we have evolved mechanisms that motivate us to behave retributively, even when the costs to our self-interest are high.

---

elsewhere—Sommers (in preparation)—I refer to the position that results from my conclusion as "metaskepticism about moral responsibility."

[4] See Elster (1990) for a careful discussion of the rationality of revenge. My point here is simply that it is often not in the individual's short term self-interest to retaliate after an offense. It may not even be in the individual's long term self-interest to retaliate (if we identify 'self-interest' as something other than 'conducive to biological fitness').

Retributive emotions or attitudes serve as these mechanisms. Trivers (1971) argues that for reciprocal altruism to emerge in hominids a feeling he called 'moralistic aggression' would emerge as a means of motivating violent retaliation against defectors. Frank (1988) refers to retributive emotions as 'commitment devices': psychological mechanisms designed to offset the desire to act in our immediate self-interest (not retaliating) in favor of long term adaptive gains. An attitude like outrage, according to Frank, serves as a tax for not behaving retributively. (If we do not retaliate we have to go on feeling outraged.) More recently, Boyd et al (2005) have proposed that a phenomenon they call 'altruistic punishment' is required for cooperation strategies to be adaptive in large groups. And Fehr and Gachter (2000, 2002) have conducted a series of experiments which show, first, that cooperators willingly suffer costs in order to punish defectors in public goods games even when they know that they will never interact with the defectors again, and second, that negative emotions like outrage or resentment motivate this behavior.

There is good reason to think, then, that retributive emotions have been selected for their contribution to social coordination and hence to fitness. But there is also reason to think that these emotions, while perhaps necessary, are not always sufficient to motivate the appropriate retributive behavior in humans. The argument for this claim is as follows.[5] Human beings are creatures of high cognitive sophistication: unlike our early hominid and primate ancestors, we can question the rationality of our emotions. By this I mean that we can question (1) whether or not a particular emotion and the accompanying behavior serves our short term material self-interests, and (2) whether or not the emotion makes sense, i.e. whether it is consistent with other beliefs we hold to be true. If we find an emotion to be irrational in either sense (but especially

[5] See Sommers (2007b) for an elaboration of this argument.

the first sense), we are motivated to resist performing the act that tends to accompany the emotion. To be sure, we are not always successful in the attempt. But it is plausible to think that the link between emotion and accompanying behavior is undermined *to some degree* by this cognitive sophistication. Rational self-interest erodes this link and prevents the retributive attitude from doing its adaptive work.

This creates an adaptive problem for cognitively sophisticated creatures like human beings. Retributive behavior is still fitness enhancing for its role in improving social coordination; but increased cognitive sophistication—which emerged for other reasons—makes it less likely that the individual will behave retributively. Human beings, then, need (or needed) something else to offset the dampening effect of greater rational capacities. This 'something else,' I argue, has taken the form of the two related but distinct sets of norms: norms about honor and norms about moral responsibility.[6]

4. Two types of cultures.

I would like to focus on two cultural environments in which the retributive emotions and behavior have important adaptive functions. The first culture has the following features:

---

[6] The term 'norms' in this paper refers broadly to social rules that regulate behavior within a group or culture. In addition, norms are *intrinsically motivating* (Sripada and Stich (forthcoming), Kelly and Stich (forthcoming)), meaning that the motivation to comply with norms does not depend solely on the perceived consequences of failing to do so. I would like to remain neutral as possible on the kinds of cognitive architecture that might underlie norm acquisition (but see Kelly and Stich (forthcoming), Sripada and Stich (forthcoming), and Nichols (2004) for excellent overviews), except to say that norms must involve more than brute emotional responses to behavior, and that norms may involve, or give rise to, *beliefs* about the kinds of behavior that are obligatory or impermissible.

(1) Cooperation usually occurs within tight-knit groups and among kin. There is relatively little cooperation among strangers.

(2) Protection of scarce resources is crucial to survival. One occasion of theft can result in loss of the entire wealth of a family or individual.

(3) Attempts to raid or steal the property of others are common.

(4) There is little or no protection from the State—the culture is relatively lawless.

Nisbett and Cohen (1996) attribute these features to herding and frontier societies. But, as they note, these characteristics may be present in inner city gang life, some tribal societies, outlaw and mafia cultures, and other types of environments as well. Following Nisbett and Cohen, let us call these types of cultures 'honor cultures.'

The second culture has these features.

(1) The economy relies heavily on cooperation and features anonymous interactions and cooperative ventures among relative strangers.

(2) The rewards of raids or thefts are not very high, and so protection of recourses is not of paramount importance. A single theft cannot cause an individual or family to lose their entire wealth.

(3) Discouraging free-riding is important for the interests of the group, not just for particular individuals.

(4) There exists some kind of governing body that polices and enforces norms about cooperation and criminal behavior.

Nisbett and Cohen claim that this type of society is realized in agriculturally based communities among others, and may be found today in relatively affluent and stable societies. Let us call cultures with these features 'institutionalized cultures.'[7]

Retributive feelings, beliefs, and behavior have important functions in both societies. But the functions are not identical. In honor cultures, the retributive dispositions serve primarily as a deterrent to theft or offenses directed at the individual and his family. If you are known as someone who is outraged at the slightest insult—a hair-trigger retaliator who will risk life and limb to punish even the mildest transgressions—then potential offenders are far less likely to take advantage of you. True, acting according to this disposition will come at great risk, but the risk is outweighed by the future prospect of having your property, wealth, and livelihood destroyed by a single offense.

In addition, since there is no effective centralized form of law-enforcement, it is crucial that the offended party be the one who does the retaliating. Third-party punishment is almost worthless. If third-party punishment does occur, the victim can no longer retaliate and so he may acquire the reputation of being an easy mark, someone whose interests may be harmed without undergoing too much risk (since third party punishment remains rare). Thus, if an individual is attacked, he must avenge the wrong himself to avoid garnering this reputation.

If these observations are correct, then we should expect the normative focus in honor cultures to be on the offended party and not the offender. And this appears to be what we find. After insults or offenses occur, there is great normative pressure on the insulted person to avenge

---

[7] The label, though not ideal, tries to draw attention to the institutions in the economy that permit anonymous cooperative endeavors and institutions that provide some degree of protection from criminal behavior. (Thanks to Alexandra Plakias for suggesting the label.)

himself. The passage by Busquet cited above describes the disadvantages and scorn heaped upon anyone who chooses not to do so. Consider too the following passage from Hasluck (1954: 231-232) about the Albanian Highlanders:

> A man slow to kill his enemy was thought "disgraced" and was described as "low class" and "bad." Among the highlanders he risked finding that other men had contemptuously come to sleep with his wife, his daughter could not marry into a "good" family. If he does, he retains his honor.

Norms about honor are extremely well suited to motivate the appropriate type of retributive behavior. Although there are significant prudential considerations that weigh against retaliation, these considerations are more than offset by the immediate fear of disgrace, ostracism, and the decrease in status for both oneself and one's family. He who does not retaliate right away becomes a target for future offenses, including "contemptuous" attempts to sleep with one's wife. And as expected, third party punishment is shunned in honor cultures. Albanians consider prison—the epitome of third party punishment—to be a "nuisance, nothing more than a delay…Prison isn't satisfying for the family." (Blumenfeld 2002: 72). Another example: Anthony James, a member of the honor culture of inner city corner life, had this to say when the man who killed his half-brother turned himself in and was charged with first degree murder: "I would have rather him stayed on the street -- and get some street justice . . . I'm very upset that I can't do nothing about it. I'm very upset that this dude took the sucker way out and turned himself in. I'm mad and angry." (Merida 2006).

In 'institutionalized cultures,' by contrast, the primary function of retribution is to discourage free-riding behavior in general. Individuals in this culture are not at risk of losing their wealth and livelihood from one raid or attempted theft, and so projecting an image of a hair-

trigger avenger is not so important. Indeed, it may be positively harmful since it will encourage unnecessarily risky attempts at revenge. In an institutionalized culture, then, we would expect the normative focus to be on the offender rather than on the offended party. What matters is that the offender be punished (to deter him and others from cheating again); not who does the punishing. Third party punishment is therefore quite useful: the identity of the retaliator is not nearly as important.

What sort of belief system would motivate the appropriate set of behaviors here? The concepts of moral responsibility and desert. If we think that offenders deserve punishment, that it would *wrong* for them not to be punished, then, again, the link between retributive feelings and the appropriate form of behavior is reinforced. Refer back to the quotation by Kant. The punishment "ought to be done in order that every one may realize the desert of his deeds, and that blood-guiltiness may not remain upon the people; for otherwise they might all be regarded as participators in the murder as a public violation of justice." There is no mention whatsoever of the offended party or his family. It makes no difference who avenges the wrong. All that matters is that the transgressor be punished, so that a great injustice does not occur throwing the universe 'out of whack.' A prison term would therefore be satisfying, as long as the offender is getting what he or she deserves.

Of course, a desire to prevent universal injustice is not nearly as strong a motivating factor as the fear of being dishonored or disgraced, but it does not need to be. Third party or altruistic punishment often comes at little cost to one's own interests. As Boyd et al (2005) have argued, altruistic punishment can evolve in large groups precisely because the punishment requires little sacrifice on the part of the punishers (as long as defections remain relatively rare). That cost is more than offset by the slight thrill of *schadenfreude* we derive from watching

defectors suffer, as well by the desire not to let injustice be committed. Men especially, according to Singer et al (2006: 467), derive pleasure from watching a defector suffer the costs of their transgression—even when they are not the target of the offense. Experimenters hooked up subjects to fMRI machines and examined their responses to watching a defector receive a shock:

> Both sexes exhibited empathy-related activation in pain-related brain areas….towards fair players. However, these empathy-related responses were significantly reduced in males when observing an unfair person receiving pain. This effect was accompanied by increased activation in reward-related areas, correlated with an expressed desire for revenge.

One might object that the feeling of schadenfreude is just that—a feeling. So where does the concept of desert come in? Note, however, that the appropriateness of the feeling in this case does seem to have some cognitive content.[8] It presupposes a belief that anyone who gets away with a crime or a norm violation *deserves* to be punished; that is how the subjects explain their satisfaction at seeing the unfair players shocked. We have this kind of feeling even when we ourselves are not the victims of the particular injustice and have no reason to expect that we will interact with the defectors in the future.[9]

---

[8] See Solomon (1976) for a classic defense of the cognitive theory of emotions.

[9] This slight thrill is in marked contrast to the ecstasy that individuals in honor cultures experience when they get revenge for a personal insult or attack. According to Djilas (1958: 107) vengeance was "our pastures and springs—more beautiful than anyone else's—our family feasts and births. It was the glow in our eyes, the flame in our cheeks, the pounding in our temples…"

Summarizing my observations thus far, honor culture environments require norms of retribution that put pressure the offended party: it is of the utmost importance for individuals to show potential future offenders that attempted thefts will come only at monumental risk to their own interests. Beliefs and norms about honor provide the necessary motivating force to retaliate against even minor offenses, often at great cost, so that this hair-trigger image is conveyed. (Boyd and Richerson's (2005) theory of gene-culture co-evolution illustrates one way for these beliefs and norms and to arise within a population. Indeed, they appeal to Nisbett and Cohen's hypothesis as a means of illustrating their model.) In institutionalized cultures, hair trigger vigilante justice is generally not worth the risk. Indeed, institutionalized cultures may feature 'turn-the-other-cheek' norms to discourage personal retaliation.[10] But it is still necessary to deter free-riding and anti-social behavior, so norms about desert and responsibility—which focus on the intrinsic "rightness" of punishing offenders—emerge to motivate individuals who may not have personally been harmed to punish at a minor cost to their own interests.

5. Some Explanatory Features of this Analysis

A better understanding of the differences between these two beliefs systems may shed light on puzzles that moral philosophers and legal theorists (mostly from non-honor cultures) have struggled with in the last century. Consider for example the phenomenon of collective punishment. In honor cultures, the targets of retaliation are often not the same people who committed the offense—something that individuals in non-honor cultures have a difficult time

[10] See Elster (1990) for a discussion of the role of these types of norms.

understanding or justifying.[11] Retributive theory in penal philosophy forbids the punishment of the innocent, but this prohibition does not appear to extend to cultures where protecting one's family's honor and reputation is paramount. Daly and Wilson (1984) write that in foraging societies revenge is not necessarily taken against the killer himself. Moreover, the preferred target must be an individual of status equal to the original victim.[12] They cite Balickci's (1970: 184) description of an Inuit tribe:

> The objective of the revenge party was not just to kill the original murderer but members of his kindred as well. In a sense the members of the kindred shared responsibility of the murder.

'Shared responsibility' is one way of interpreting this behavior. But it may also be a way for someone brought up in a culture which emphasizes moral responsibility norms to explain the behavior of individuals who do not see the world in those terms. Perhaps a more plausible interpretation is summed up by Clint Eastwood's line in the film *Unforgiven* just before he kills a protesting Gene Hackman: "deserve's got nothing to do with it." In other words, honor cultures focus less on the whole concept of responsibility and, in particular, the 'deservingness' of the offender. For the Montenegrins, "revenge was taken in behalf of the soul of the man who was being avenged, *rather than to punish the specific killer*…vengeance symbolically replaced to the victim's clan the blood that had been lost" (Boehm 1985: 60, my italics). The important thing is that the injured party retaliates against *someone*, someone who bears a connection to the offender. Otherwise, honor is lost. Of course, the most suitable target is the murderer (as long

---

[11] See Feinberg (1969, 1974) for classic analyses. See also the 2006 issue of *Midwest Studies,* the entire issue of which is devoted to collective responsibility.

[12] Daly, M. and Wilson, M., *Homicide.* (Aldine de Gruyter, 1988)

as he is of equal status to the victim). But there is no prohibition against punishing relatives or associates of the offender, since the primary function of retaliation is to restore the reputation of the offended party. The result is that individuals in honor cultures may have radically different ethical perspectives about justice, courage, and the appropriate targets of punishment—perspectives that the language of moral responsibility cannot capture.

This analysis may also shed some light on the terrible and almost incomprehensible practice of honor killings.[13] This is when a family member murders a woman or girl because she has lost her virginity, which is seen as a stain on the family's honor. What makes these cases even more unfathomable is that often the murdered woman is a victim of rape. Most groups in Western cultures find this perspective not just appalling but utterly bewildering. Even if one believes that premarital sex is a mortal crime, how can anyone think it is appropriate to punish victims of rape? It is clearly not the rape victim's fault that she committed this "crime" and lost her virginity. The phenomenon becomes a little easier to understand if we imagine that the cultures that support this behavior do not have a strong notion of desert to begin with. If that is the case, it almost irrelevant that the victims do not in any way deserve their fate. Note that in many cases, the family members are not moral monsters; they are often consumed with grief at the death of a *beloved* daughter or sister. But they will seldom repent the act. We might imagine that family members regard victims of rape as a fatal illness.[14] That is why tremendous grief is consistent with a complete lack of repentance for the killings. The *injustice* of these deaths does

---

[13] Sommers (in preparation) presents a more detailed analysis of the connect between honor killings and responsibility.

[14] According to one tribal leader, "a woman is like an olive tree. When its branch catches a woodworm, it has to be chopped off so that society stays clean and pure." (Feldman 2001).

not resonate to members of honor cultures nearly as much as it does to those of us in institutionalized cultures.

Finally, this analysis may apply not only to cultures with individuals whose interests are especially vulnerable to attack, but to entire countries or cultures that face this predicament as well. If so, the analysis may help to explain one curious feature of Israeli society—an almost obsessive aversion to being labeled as "freiers."[15] From the time they are children Israelis learn that being called a freier—roughly, 'chump' or 'sucker—is the worst insult of all. Indeed, it amounts to loss of honor. The 'sucker syndrome' or 'freier factor,' as it has been called by the Jerusalem Post, permeates every aspect of society—from the Government's role in the peace process,[16] to sexual relationships, to the way Israelis drive (to 'drive defensively' is to be a freier and as anyone who has visited Israel knows, there is not a lot of defensive driving on the highways). This phenomenon—some call it an 'illness,' others speak of it with pride—has, by general consensus, its roots in the two thousand years Jews have spent as victims, culminating in near annihilation in the holocaust. Moreover, since its inception in 1948 Israel has been in a precarious situation, one war away from ceasing to exist. The plight of Israel is Nisbett and Cohen's frontier-herding society writ-large, and so it's no surprise that honor plays a large role in Israeli culture and even the tiniest indications that one is being taken advantage of are shunned at all costs.

6. Caveats

[15] See Bloch, L. (2003) for an interesting discussion of the freier phenomenon.

[16] In 1998, Benjamin Netanyahu broke off talks with Albright and Dennis Ross, saying "we are not freiers—Israel cannot give and give and not get anything back in return."

Some caveats are certainly in order. First, although at times I have presented two types of retribution motivators as distinct, it is more plausible to see them as opposing ends of a continuum. My claim is that societies that have features in common with 'honor cultures' will place greater emphasis on honor and how the *offended party* ought to act. Societies that have more in common with institutionalized cultures will place more emphasis on moral responsibility and how the *offender ought to be treated*. Naturally, there will be overlap—a notion of honor in institutionalized cultures, and a notion of moral responsibility in honor cultures. (It would be interesting to examine whether the notion of moral desert is more prominent in close interpersonal relationships in honor cultures.[17]) But if I am correct there should be an identifiable difference in emphasis in reactions to similar offenses.

Second, there will certainly be some within-group variation which may be attributed either to learned or innate differences. One fascinating example is Laura Blumenfeld, a Washington Post journalist and author of a book called *Revenge*. Blumenfeld, a quiet diminutive Harvard graduate from a Jewish middle class family in Long Island embarked on a decade-long quest for revenge after her father was shot as a tourist in Israel. Since her father suffered no long term injury, and the shooter was in prison, this obsession with revenge baffled her friends and family (including her father, the victim, who was completely satisfied that justice had been carried out). When asked by her friend Rachel for an explanation for her obsession, Blumenfeld was almost too embarrassed to reply. After making Rachel promise not to laugh, Blumenfeld replies: "I wanted to them to know…"

"That you can't shoot your Dad?"

---

[17] Thanks to Joshua Knobe, John Doris, and Shaun Nichols for suggesting this possibility.

"No, not that…that you can't…*you can't fuck with the Blumenfelds!...*That there's someone you have to answer to…" (Blumenfeld 1992: 114)

Blumenfeld alone (her father and brother did not participate in her quest) felt her family had been dishonored by the attack and the prison-term was in no way a remedy for the dishonor. This is exactly the type of 'honor' norm that I have been describing—one that focuses on the victim and his family.[18] But it comes from a most unlikely source.

7. Testing the Hypothesis

One might think that this account is merely a farfetched piece of speculative anthropology—a prototypical 'just-so story.' Fortunately, there are ways of testing this hypothesis using methods currently employed in the field. Henrich et al (2004), for example, assemble a series of ethnographic studies that test for punitive behavior in small scale societies. Field workers perform a series of experiments involving ultimatum games, dictator games, and public goods games on subjects from fifteen diverse cultures. This kind of experiment would provide an excellent test for my view that desert plays less of a role in honor cultures. One prediction is that altruistic punishment in public goods games would occur less frequently in honor cultures than in non-honor cultures, especially when the punishment is effected

[18] Blumenfeld's ultimate means of revenge does in part focus on the offender—she wants the shooter to recognize that her father was a human being and not merely an instrument for making a political statement (a notion that appears to be in line with Jean Hampton's justification of the "retributive idea" (Murphy and Hampton (1988).) When her mother mentions that she seems less interested in revenge, however, Blumenfeld replies: "I still want to stand up and say "don't fuck with the Blumenfelds." In the end, Blumenfeld appeals to an interesting (and rather rare) blend of both honor and desert-based norms in order to explain her behavior.

anonymously. In public goods experiments, the cost of the defection is distributed to the rest of the group and thus the defection is less likely to be seen as a personal insult. Furthermore, an anonymous punishment—at a cost to oneself with no hope of regaining the cost in the future—cannot aid one's reputation. Thus there is a little dishonor in refraining since no one will observe the individual's behavior no matter what he does, and in any case there was not much of a personal insult to begin with. On the other hand, one might expect to see unfair offers in non-anonymous ultimatum games rejected at a much higher rate in honor cultures, because in ultimatum games the insult is directed personally at the individual.[19] One might even expect to see rejection of low *randomly generated* offers from a computer rejected in ultimatum games, because of the importance of establishing reputation.[20]

For institutionalized cultures, on the other hand, altruistic punishment in public goods experiments should occur more frequently. When beliefs about responsibility and desert are the guiding norms, then it is not as important that the offense was directed at the punisher. What matters is that the defector not get away with his behavior: he *deserves* to be punished, and a public goods game allows the participants to effect that punishment at a minor cost. There should also be frequent rejections of unfair offers ultimatum games, as again, the defector deserves to be punished for his behavior. (But perhaps at not such a high cost, since the insult will not be taken quite as personally.)

---

[19] This experiment would be complicated because of other factors that enter into decision-making in ultimatum games. In some cultures, individuals abhor being in another individual's debt (for fear of being asked to reciprocate the favor) and so reject offers that exceed the fair .

[20] Thanks to James Woodward for suggesting this experiment.

Since the experiments in Henrich et al (2004) are not designed to test this particular hypothesis, it is difficult to interpret their results for my analysis. The primary purpose of this section, however, is to show that this account about the de-emphasis of desert norms in honor cultures is not an unfalsifiable just-so story—it can indeed be tested, and with methods similar to those that are presently being employed.[21] One very helpful addition would be to perform interviews after the experiments, asking the subjects why they behaved as they did.

8. Conclusion: Philosophical and Social Implications

If my analysis of the difference between honor and non-honor cultures is correct, or even on the right track, then there are some interesting and perhaps urgent implications that ought to be explored. Consider, for example, what a thorough examination of honor cultures might mean for current policies in Iraq. Former Army Major William McCallister has claimed that for most Iraqis the concepts of shame and honor are as important as land or water. These concepts form the moral currency of the country. McCallister (2005) attributes our unpopularity with Iraqi tribes in large part to our failure to understand this feature of their society and employ it to our advantage. Similar observations have been made about all the countries in that region, including, as I have noted, Israel. It seems plausible that in order to shape our policies properly, we need a comprehensive analysis of the role that honor plays in these countries—and the role that individual moral responsibility may not play. Once we are alerted to the differences in ethical

[21] In an unpublished study, Haidt et al (under revision) showed subjects scenes from movies designed to elicit retributive emotions and desires. The subjects were then asked to rate alternative endings for how satisfying they were. An experiment like this designed for cross-cultural analysis would also provide an excellent test for my hypothesis.

perspectives, we will no longer expect people in honor cultures to immediately embrace our own notions of justice. We will have a much more complete understanding of the moral motivations of their citizens.

The same line of thinking applies to domestic policies. Many criminal groups—inner-city gangs and the mafia for example—are thought to be honor cultures. A deeper understanding of the codes and norms of these groups will undoubtedly help address the root causes of criminal behavior. It will also raise some deep philosophical issues in the justification of punishment. It is likely that our notions of positive retribution (the view that the culpable *should* be punished, independent of any consequentialist benefits) and negative retribution (the view that *only* the culpable should be punished) are not as universally held as is commonly supposed. These retributive concepts will likely strike an intuitive chord for members of institutionalized societies, but will be far less intuitive in cultures where honor is the primary motivator for revenge behavior. Criminal justice in America is in part grounded in retributive principles that involve a strong notion of moral desert.[22] It may be that we are punishing criminals according to principles that (a) cannot be universally justified, and (b) are not shared by individuals in many groups throughout the United States.

Finally, for similar reasons, this analysis raises tough challenges for libertarian and compatibilist accounts of moral responsibility which posit certain universal necessary and/or

[22] See Duff (1996) for an overview of contemporary theories of punishment. Duff notes that the 'retributivist revival' in the 1970's stemmed in part from the perceived moral failings of consequentialist justifications. Retributivist theories attempt to provide an 'intrinsically appropriate' response to crime—but the response can only be directed at the offender.

sufficient conditions for genuine blameworthiness or praiseworthiness.[23] Individuals in honor cultures will likely not find it remotely plausible that only agents who acted with contra-causal free will, or according to reason, deserve blame or punishment. (Indeed, at times the agent does not need to have *acted at all* in order to be a fair and appropriate target of punishment, as illustrated in the remarks by Daly and Wilson.) Consequentialist versions of compatibilism will have similar problems as there is a strong deontological streak in the honor-culture vengeance. Individuals in honor cultures often believe they have a 'sacred duty' to avenge an offense, one that may have nothing to do with any consequentialist considerations. The individual is duty-bound to retaliate not because the offender was reason-responsive, or had a strong metaphysical ability to do otherwise, or even out of a conscious belief that the retaliation will benefit the individual in the future, but rather because it is simply disgraceful to allow an attack to be unavenged, no matter what state of mind the offender was in.

One might object that the goal of compatibilist and libertarian theories is not to capture the intuitions of Jibaro Indians or Montenegrin tribesmen but rather to give a *true* account of the conditions of moral responsibility. After all, the Jibaro Indians likely have some false views about the origin of the Earth as well, but this does not undermine well-founded geological theories. But this objection fails to appreciate the role of intuitions in evaluating the soundness of philosophical arguments about free will and moral responsibility. All contemporary theories

---

[23] Some necessary conditions that have been proposed in libertarian theories of responsibility are: (1) the agent must have a strong metaphysical ability to do otherwise and (2) the agent must have self-caused some part of his or her character. Some necessary conditions in compatibilist theories are: (1) the agent must have acted from a second order desire; and (2) the agent must be responsive to reasons. Of course, this list is not exhaustive.

of moral responsibility appeal to intuition for the justification of key premises, cases, and principles (van Inwagen's (1983) famous beta-principle is one example.[24]) If my hypothesis is correct, however, individuals in honor cultures have *fundamentally different* intuitions about the appropriate targets of blame and punishment. While these intuitions may be based on irrational or false beliefs in some cases, it would be hard to argue that this is true in all cases.

Another way to look at this argument against universal theories of moral responsibility is to compare it to J.L Mackie's argument from relativity.[25] Concerning disagreement about moral values, Mackie (1977: 37) writes:

> The argument from relativity has some force simply because the actual variation in the moral codes are more readily explained by the hypothesis that they reflect ways of life than by the hypothesis that they express perceptions, most of them seriously inadequate and badly distorted, of objective values.

---

[24] Van Inwagen himself admits that intuition plays a key role in justifying Beta: "I must confess that my belief in the validity of Beta has only two sources, one incommunicable and the other inconclusive. The former source is what philosophers are pleased to call "intuition".... The latter source is the fact that I can think of no instances of Beta that have, or could possibly have, true premises and a false conclusion." (Van Inwagen 1983: 98-99). Van Inwagen here is referring to the beta rule regarding free will, but in his discussion of the 'direct argument' he explicitly applies it moral responsibility as well. Note that his two sources for believing beta to be valid are really quite similar, since someone with radically different intuitions about responsibility could presumably come up with counterexamples rather easily. Sommers (forthcoming) discusses the role of intuitions in the free will debate at greater length.

[25] Thanks to Don Loeb for making the analogy to Mackie's argument explicit in personal communication. See Loeb (1998) and Doris and Plakias (forthcoming) for excellent discussions of the argument from disagreement in the metaethical debate.

I believe this strategy is even stronger when applied to theories moral responsibility, for in this case the primary objection to the argument from relativity will hardly get off the ground. The 'well known counter' (as Mackie calls it) runs as follows: while we may find disagreement about a wide range moral issues, we also find widespread cross-cultural *agreement* about certain core ethical principles. These principles, the objection continues, may then be applied in different ways depending on the culture; but the principles themselves—the 'don't harm unnecessarily' principle, for example—are universally accepted. This objection might seem plausible regarding some moral values, but if my observations are correct, there is simply no widespread cross-cultural agreement about core principles of moral responsibility. Honor cultures have a diminished notion of deservingness, and what appears to be a vastly stripped down control condition, which institutionalized cultures consider to be the very essence of moral responsibility. And what principle could be more fundamental to Western notions of moral responsibility than the claim that only the culpable, only those who had control over their actions, deserve blame and punishment?[26]

Of course, it is possible that honor cultures are simply wrong about moral responsibility, or that they are not in a position to perceive the true conditions of justified blame and

---

[26] Retaliation and honor killings in honor cultures have similarities to a 'strict liability' theory of punishment, one where a control condition for responsibility is virtually absent. Non-honor cultures consider strict liability approaches fundamentally unfair, although it is invoked under certain occasions for consequentialist reasons. As Nagel (1979: 31) puts it: 'strict liability may have its legal purposes, but seems irrational as a moral position." But perhaps denying the control condition only "seems irrational" to people in certain kinds of cultures. Retaliation towards people who had nothing to do with the original offense appears to be *completely justifiable* (as opposed to 'regrettable but necessary') in the minds of individuals in honor cultures.

punishment. And it is possible that these differences in considered judgment might be resolved with sufficient conversation, philosophical analysis, and reflection. Perhaps “fully informed” people, stripped of every irrational bias and religious conviction, would converge on the same judgments regarding moral responsibility. The evidence suggests, however, that existing variation is a natural result of the different cultural and social environments in which these intuitions and judgments emerge. If honor norms are more suitable for a certain type of environment then there is nothing irrational about placing more emphasis on honor than individual moral responsibility *in those environments*. The intricately laid out conditions for true blameworthiness that arise from intuitions formed in *our* cultural environment, then, would simply have no place in honor cultures. Theories which contain conditions or criteria for justified assignments of moral responsibility that are meant to apply *across cultures* would therefore be undermined.[27]

I concede that these arguments must be developed more carefully and that the evidence in favor of this view is far from decisive. (Sommers (in preparation) presents more detailed analysis of how cross-cultural differences undermine philosophical theories of moral responsibility, punishment, and desert.) The claim that philosophical analysis and ‘wide reflective equilibrium’ cannot resolve cross-cultural disagreement about moral responsibility is in large part an empirical one. Indeed, one aim of this analysis is to illustrate the vital importance of empirical inquiry to the philosophical debate about moral responsibility. If the claim is true, however, the implications are significant. Instead of viewing our attitudes, beliefs, and intuitions about blame and punishment as tracking some kind of universal moral truth (to be formalized subsequently

---

[27] I do not see this as a pessimistic conclusion. Sommers (2007a) argues that skepticism about moral responsibility would not adversely affect interpersonal relationships, moral worth, and ethical behavior.

into a universal theory), we should regard them as responses to the various features of different environments. Some beliefs might be more or less rational within a particular environment, but no theory can describe the truth about responsibility in all environments. The analysis developed in this paper is sufficient, I believe, to make this conclusion a genuine possibility. At the very least, a defender of a *universal* condition-based concept of moral responsibility will have to answer the following question: why should the notions of moral responsibility and desert have any more objective or metaphysical status than the notions of honor and dishonor? If we are to say, with Kant, that it is an objective truth that a criminal *deserves* to suffer for his crime, then why should we not say that a family is *truly dishonored*, again in an objective sense, when they do not avenge a wrong committed by another tribe or individual?

Tamler Sommers

University of Minnesota, Morris

REFERENCES

Balikci, A. 1970. *The Netsilik Eskimo.* The Natural History Press

Bloch, L. “Who's afraid of being a freier: The analysis of communication through a key cultural frame” *Communication Theory*, 13 (2003) 2: 125-159

Blumenfeld, L. 2002 *Revenge: A Story of Hope.* Washington Square Press.

Boehm, C., 1985. *Blood Revenge.* (University of Pennsylvania Press,

Boyd, R. et al, 2005. “The Evolution of Altruistic Punishment.” In *Moral Sentiments and Material Interests*, ed. Gintis,H. Bowles, S., Boyd, R. and Fehr, E. MIT Press.

Busquet, J.1920. *Le droit de vendetta et les pacii corse.* Pedone

Daly, M. and Wilson, M., 1988. *Homicide.* Aldine de Gruyter.

Djilas, M. 1958. *Land Without Justice: An Autobiography of His Youth.* London: Methuen.

Doris, J. M., and Plakias, A. forthcoming. “How to Argue About Disagreement: Evaluative Diversity and Moral Realism.” In W. Sinnot-Armstrong (ed.), *The Biology and Psychology of Morality*. New York: Oxford University Press.

Duff, R.A. 1996. “Penal Communications: Recent Work in the Philosophy of Punishment. *Crime and Justice,* 20: 1-97

Elster, J.1990. “Norms of Revenge,” *Ethics.* 100: 862-895

Fehr, E. and Gachter, S. 2002. “Altruistic Punishment in Humans.” *Nature*. 415: 137-140.

Fehr, E. and Gachter, S. 2000. “Cooperation and Punishment in Public Goods Experiments,” *American Economic Review* 90: 980-994;

Feldman, I. 2001. “All in the Family.” *The Review.* January, 2001.

Feinberg, J. 1974. *Doing and Deserving.* Princeton University Press.

Feinberg, J. 1968. “Collective Responsibility.” *Journal of Philosoph*y. Nov. pp. 674-688.

Frank, R. 1988. *Passions Within Reason*. Norton.

Haidt, J., Sabini, J., and Worthington, E. (n.d.). “What exactly makes revenge sweet?” (Manuscript under revision)

Hasluck, M. 1954. *The Unwritten Law in Albania.* Cambridge: Cambridge University Press. pp. 231-232

Henrich, J. et al. 2004. *Foundations of Human Sociality.* Oxford: Oxford University Press.

Kant, I. 1996. *The Metaphysics of Morals.* Cambridge: Cambridge University Press.

Kelly, D and Stich, S. forthcoming. “Two Theories About the Cognitive Architecture Underlying Morality, to appear in P. Carruthers, S. Laurence & S. Stich, eds., *The Innate Mind*, Vol. III, *Foundations and the Future*.

Knobe, J. and Doris, J. forthcoming. “Strawsonian Variations: Folk Morality and the Search for a Unified Theory.” to appear in J. Doris et al. *The Handbook of Moral Psychology*. Oxford: Oxford University Press

Loeb, D. 1998. "Moral Realism and the Argument from Disagreement," *Philosophical Studies*, 90: 281-303.

Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. Penguin.

Machery, E. et al. 2004. “Semantics, Cross-cultural Style.” *Cognition*, 92, 3, B1-B12.

McCallister, W.S.2005. "The Iraq insurgency: Anatomy of a tribal rebellion," *First Monday,* 10:3.

Merida, K. 2006. "In or Out of the Game?" *Washington Post.* December 31, 2006.

Murphy, J. and Hampton, J. 1998. *Forgiveness and Mercy*. Cambridge University Press.

Nagel, T. 1979. "Moral Luck" in *Mortal Questions*, Cambridge University Press.

Nichols, S. 2004. *Sentimental Rules*. Cambridge University Press.

Nichols, S., Stich, S., and Weinberg, J. 2003. "Metaskepticism: Meditations in Ethno-Epistemology." In *The Skeptics,* ed. S. Luper. Burlington, VT: Ashgate, 227-247.

Nisbett, R. and Cohen, D. *Culture of Honor*. Westview Press.

Richerson, P., and Boyd, R. 2005. *Not by Genes Alone.* University of Chicago Press.

Singer et al. 2006. "Empathic Neural Responses are Modulated by the Perceived Fairness of Others." *Nature:* 439: 466-469.

Solomon, R. 1976. *The Passions.* Anchor/Doubleday.

Sommers, T. forthcoming. "More Work for Hard Incompatibilism." *Philosophy and Phenomenological Research.*

Sommers, T. in preparation. *Relative Justice*. Book manuscript. Under contract. Princeton University Press.

Sommers, T. 2007a. "The Objective Attitude." *The Philosophical Quarterly.* 57, 228: 321-342.

Sommers, T. 2007b. "The Illusion of Freedom Evolves," in *Distributed Cognition and the Will*. Spurrett, D., Kincaid, H. Ross, D., Stephens. L. eds. MIT Press.

Sripada, C and Stich, S. "A Framework for the Psychology of Norms," in P. Carruthers, S. Laurence & S. Stich, eds., *The Innate Mind: Culture and Cognition*. Oxford University Press, 2006. Pp. 280-301.

Trivers, R. 1971. "The Evolution of Reciprocal Altruism." *Quarterly Review of Biology*, 46:35-56.

Van Inwagen, P, 1983. *An Essay on Free Will*. Oxford: Clarenden Press