

[in press, *Intellectica*, June, 2006]

Running Head: Mental Representation and Natural Selection

Mental Representation and Natural Selection: The Special Case of Human Social Evolution

Jesse M. Bering

Department of Psychology

University of Arkansas

Todd K. Shackelford

Department of Psychology

Florida Atlantic University

This article draws from and expands some of the arguments presented in Bering and Shackelford (2004). Address correspondence to Jesse M. Bering, Institute of Cognition and Culture, Queen's University, Belfast, Northern Ireland BT7 1NN. Tel: +44 (0) 28 9097 3706. Fax: +44 (0) 28 9097 3700. (email: j.bering@qub.ac.uk).

Abstract

In this article, we argue that (1) there are several “Darwinian algorithms” that are specific to humans because they are grounded in a novel representational system allowing for abstract causal reasoning in the social domain and that (2) were it not for the problems caused by this mental representational capacity, the behavioral adaptations spawned by these algorithms would not have evolved. Such behavioral adaptations may include *information-retention homicide*, the killing of others who are believed to possess information about the self which, if publicly disclosed, has the potential to reduce reproductive potential or inclusive fitness, and also *confession* and *suicide*, which may necessitate the capacity for self-referential emotions, such as shame. Human psychology, unlike the psychological systems of other closely related species (such as chimpanzees), is characterized by a mental representational system that presented adaptive problems fundamentally unique to our species.

## Mental Representation and Natural Selection: The Special Case of Human Social Evolution

In several ways, mainstream evolutionary psychology is similar to the behaviorism of the mid-20<sup>th</sup> century. This is ironic, considering one of behaviorism's most infamous claims was that differences in the psychological systems between species were negligible because there exists only a single domain-general learning mechanism. This argument is antithetical to the position held by evolutionary psychologists, who have empirically demonstrated specialized psychological functions in human beings that are dedicated to solving specific problems in the species' recurrent environments. But consider the following. Both behaviorism and evolutionary psychology are primarily concerned with how observable behaviors are affected by local environments over time; both fields do not attend to the influence of intentional motives, and consciousness in general, arguing that mental states often obscure true causal processes, and; by neglecting the role of consciousness, both fields treat the emergence of human behavior no differently than they treat the emergence of behaviors in other species.

Evolutionary psychology is unlikely to suffer the same theoretical fate as the now defunct behaviorism of old. Its explanatory relevance is significant and decisive, providing a unifying meta-theoretical framework within which all of psychological science can be organized (Buss, 1999). However, to provide a comprehensive account of human behavior, as well as cultural variability, evolutionary psychology must begin to address how one particular cognitive system, that allowing for thoughts about others' thoughts, had on the social evolution of our species. If it is truly to become the new science of the *human* mind, researchers in this field must evaluate how behavioral categories that are unique to humans, such as suicide and some categories of murder (as well as potentially many other social behaviors), came about through selective pressures that were also unique to humans. In this article, we argue that the evolution of a capacity to entertain others' mental states—a “theory of mind”—forged a new path in natural selection, one where behaviors designed to capitalize on social reputation were strongly favored. Furthermore, we show how ancient behaviors that are embedded in our primate ancestry were influenced or changed by the evolution of this distinctive system.

To date, evolutionary psychology has mostly paid only lip service to this psychological system, usually referring to it as a theory of mind “module” used to generate explanations and predictions for other people’s behavior and therefore as adaptive in ancestral social environments. Although this is accurate, it does not penetrate the issue deeply enough. Throughout hominid evolution, there were likely many pathways of adaptive processes linked to the presence of a theory of mind, and any account of such pathways must sufficiently explain how new human adaptations were *causally determined* by this system. Thus, we must take into consideration several factors, including: (1) which preexisting primate adaptations were meaningfully affected by uniquely human social cognition; (2) the precise means by which these preexisting adaptations were influenced or modified by the presence of a theory of mind; and (3) the reproductive payoffs of fundamentally novel behaviors made consequent by these coactions between the old and new (see Povinelli, Bering, & Giambrone, 2000).

Any fitness-enhancing psychological adaptations that built on these coactions can therefore be said to have required this unique feature of the human mind to evolve. (Another way to say this is that the preexisting adaptations that had operated independently of theory of mind were necessary but not sufficient for human evolution to unfold the way it did). Because new adaptations are constrained by features of preexisting structures (Buss et al. 1998; Dawkins, 1986; Dennett, 1995; Williams, 1992), selection favors traits that can exploit existing designs without interfering with their functioning (see Buss et al., 1998; Gigerenzer, 2001). We believe that because the new adaptive problems caused by theory of mind were so substantial, however, some of these preexisting designs had to be fundamentally revamped. In certain cases, what was behaviorally *adaptive* before theory of mind evolved became behaviorally *maladaptive* after it became canalized in human brains.

In the present article, we offer concrete examples of three behavioral categories where theory of mind would have had its most obvious effects on the evolution of human social behavior. We focus our attention on confession, certain types of homicide, and suicide. We show how each of these behaviors operates via standard mechanisms of inclusive fitness but, in addition to this, appear to be adaptive solutions to the complex mental representational contingencies of human social cognition.

*How Theory of Mind Changed the Course of Human Social Evolution*

Primate evolution is characterized by increased social complexity, with specific adaptations emerging to process information from many different social domains. Recurrent exposure to information within each of these domains has resulted in cognitive adaptations in different primate species specialized for producing ancestrally-adaptive behavior. Although species-atypical experiences can lead to the development of such adaptive behaviors astray and suggests considerable neural plasticity (an integral part of selective processes, see Bjorklund, 1997; Bjorklund & Pellegrini, 2002), throughout evolutionary history organisms were faced with species-typical experiences that served to carve out psychological mechanisms designed to respond to particular environmental factors.

In human societies, which consist of fluid polities where one's social position is never stable, variables correlated with aggression and violence, and often implicated in cases of homicide, include status- and reputation-related threats (Daly & Wilson, 1988). Because increased status is linked to greater acquisition of resources that facilitate greater reproductive opportunities (e.g., higher-status males are likely to have economic incentives that attract the sexual interest of females), assaults to one's reputation and status are likely to motivate feelings of hostility and vengeance that may be channeled into aggression. Although female-female physical violence is less frequent in humans, when it does occur, it is often precipitated by one individual seeking to undermine the reproductive opportunities of another by spreading rumors about her sexual infidelity or her otherwise "loose" behavior with males (Ahmad & Smith, 1994; Buss & Dedden, 1990; Campbell, 2002). By physically assailing the source of these rumors, the aggressor may be reducing the likelihood of such insults occurring in the future. Similarly, among males, barroom brawls and schoolyard fights are often precipitated by personal sleights, homophobic remarks, and insulting comments regarding relatives, each of which is designed to affect dominant and subordinate relations (Ahmad & Smith, 1994; Atzwanger, 1995; Buss & Dedden, 1990; Daly & Wilson, 1988; Gladue, 1991).

The importance of language in directing these status-related threats is deliberately underscored in the previous examples. Along with its capacity to threaten one's standing in a social community, the evolution of language provided individuals with social information that could not be inferred through first-person access to behaviors (Dunbar, 1993). In the absence of language, selfish acts would have had enormous payoff in

interactions with subordinate others, just as they appear to do so for nonhuman social primates. This is because, as Dunbar (2004:102) points out, “monkeys and apes are constrained in what they can know about their [social] networks. They know only what they see.” In contrast, among humans, impulsive acts of unguarded selfishness, observed not just by dominant others but instead by almost any person with a tongue, are frequently maladaptive because of potential punishment rendered by *absent* third-parties. Social information had the potential to grasshopper between minds *ad infinitum*.

Because phylogenetically older social psychological adaptations evolved prior to the unique mental representational features of human social cognition, however, and therefore did not evolve to be sensitive to the epistemic states of others, human psychology seems to continue to grapple with strong heuristics motivating (previously adaptive) selfish behavior, such as violence and sexual coercion in males (e.g. Daly & Wilson 1988; Thornhill & Palmer 2000). Hence Goethe’s self observation, “*There is no crime of which I do not deem myself capable.*” During the course of human evolution, the advent of language and a representational theory of mind meant that social information was no longer confined to making *in situ* observations of others’ behaviors, nor temporally constrained, but could now be rapidly communicated through second-hand sources and across large expanses of distance and time, laying the groundwork for complex mechanisms of social justice and organized regimens of third-party punishment (punishment of norm violations by “unaffected” *absent* third parties; Fehr & Fischbacher 2004).

What is unusual about human sociality is the fact that third-parties who are physically *absent* during social encounters still mediate an actor’s decision-making toward the second-party, normally by gearing the former toward altruistic behaviors or norm compliance. This is because, due to their theory of mind, subordinate second-parties are able to mentally represent an absent third-party’s state of ignorance about the unobserved event and are strategically motivated—and emotionally driven—to disclose their victimization with naïve third-parties through declarative language.

With the appearance of language, social information could be rapidly dispersed and the amount of information potentially harbored about others was significantly increased. Knowledge about specific others (“Jim said that Mary cheated on John while they were dating”) could be stored in “social data files” and retrieved when

determining one's behavior towards these others ("Therefore, if Mary asks me out on a date, I'll turn her down"). The more knowledge one has of the behavioral and personality characteristics of specific others, the more adaptive one's behaviors concerning these individuals (e.g., because Mary has expressed signs of infidelity in the past, establishing a long-term relationship with her is risky because of an increased possibility of cuckoldry). Although repeated encounters with others in a variety of situational contexts can help individuals assess the type of person someone is, the possibility of deceit threatens the reliability of these first-person character judgments. Therefore, detecting deceptive traits in others—traits that may jeopardize one's own genetic fitness (e.g., a prospective job applicant's history of tax evasion; a potential mate's questionable sexual orientation or allegations of child abuse from a previous spouse)—involve: (1) gathering social information "through the grapevine" of language (2) responding emotionally to such information, and; (3) responding adaptively to such behavior (e.g., offering the job to someone else or breaking off the engagement).

With the co-evolution of theory of mind and language, our evaluative judgments of others were no longer confined to, say, who we saw grooming whom, noticed was absent during oestrus, or observed attempting to steal food, but rather we could be informed of others' behaviors even if they occurred during our absence. Dunbar (1993) has reported that as much as 60% of conversation involves social gossiping, and other investigators have shown that people are primarily interested in the misdeeds and misfortunes of others, especially when these others are in positions of authority (McAndrew & Milenkovic, 2002). This pattern appears early in development, with school-aged children frequently participating in name-calling and rumor transmission, each of which has consequences for a child's popularity and the number of friends he or she has (Crozier & Skliopidou, 2002). Mechanisms designed to monitor or control the flow of self-related information also seem at play early in development, for example in children's denouncement of others as "blabbermouths" and "tattletales" when their misdeeds are exposed.

An obvious advantage of gathering strategic information about social others is the increased likelihood of avoiding threats to genetic fitness before they happen (see Shackelford & Buss, 1996). Having knowledge of a dating partner's history of alleged physical abuse against his ex-wife, a pattern that did not appear until three years into the marriage, can help a woman to make an informed decision if this man decides to propose to her.

Much like Pascal's wager, even if such information were unreliable, the risk associated with ignoring it is much greater than the risk of having it influence one's decision-making. Furthermore, strategic information can provide social leverage in the context of status-striving and resource acquisition, affording power over others who fear their social exposure (e.g., blackmail).

This strategy of deriving fitness-related "negative" information about others can, not surprisingly, backfire. Others' attention also homes in on our own undesirable traits and peoples' solicitation of negative information about us can pose an enormous threat to our genetic self-interests. Sensitive information may be revealed to sources that are motivated to interfere with our reproductive opportunities or otherwise diminish our inclusive fitness. The result of this evolutionary dynamic can be envisaged as an arms race between individuals coexisting in social communities where *information-gathering mechanisms* (those mechanisms designed to collect information about social others) become in conflict with *information-retention mechanisms* (those mechanisms designed to block others' access to negative information about the self).

The individual with deficient information-gathering mechanisms risks engaging in social behaviors with dire genetic consequences, and the individual with deficient information-retention mechanisms runs the risk of being excluded from forms of social courtship associated with reproductive opportunities. Failure to protect sensitive information from public exposure can have many indirect effects on one's fitness. For example, social stigmas often adhere to entire families rather than to the sole family member who provoked such public aversion through his or her faulty traits or moral transgressions. Because social dissidence is often perceived as sanguineous in nature, related individuals also are likely to bear the expense of having the self's negative attributes exposed. Evidence of such social stigma has been found among individuals with mentally ill relatives (Choi, 1996; Veltman, Cameron, & Stewart, 2002), children whose mothers are HIV positive (Murphy, Roberts, & Hoffman, 2002), children and adults whose family members were murdered or committed suicide (Clements & Burgess, 2002; Dunn & Morrish-Vidners, 1988), those whose family members committed homicide (May, 2000), adult daughters of battered women (Humphreys, 2001), family members of homosexuals or bisexuals (Hammersmith, 1987; Herek, 1998), adoptive families (Miall, 1987), and children whose fathers are incarcerated (Gabel, 1992). Such information can constitute the anatomies of a family's closeted skeletons; from an



evolutionary perspective, reluctance to make this information public occurs because the content of such secrets can have deleterious effects on reproductive potential.

*Social learning and developmental modifiability of information-retention mechanisms*

Human cultures are the aggregate output of evolved psychologies operating in concert with regular variance in socio-ecologies (see Tooby & Cosmides, 1992). Because of this natural variance, children's developing information-regulation systems must, over the course of their ontogeny, be "fine-tuned" to the particular contingencies of their environments. (This may account for why children are less likely to be harshly penalized for committing social transgressions, i.e., crimes, than adults involved in the same behavior.) A case in point is infanticide. Depending on socio-ecological conditions, infanticide is either culturally expected (resource-limited environments) or culturally demonized (resource-rich environments). In the case of the former, the information-retention mechanism would not be activated (in certain cultures, mothers feel no need to keep private the fact that they have killed a sickly twin, for example), whereas in the latter it would. Although there are likely broadly based default system features, such as being motivated to retain from social others personal information concerning disease or sexual behavior, social learning programs these default systems with specific content. Therefore, the particular cultural manifestation of the information regulatory system will be carved out through developmental processes that lend social evaluative weight to particular behavioral types. Transgressing behavioral norms should give rise to shame—a very forceful emotional trigger for information-retention mechanisms. Likewise, others' transgressions of these behavioral norms should give rise to anger, disgust, and vengeance. This position does not question the existence of "deep structures" or "moral syntax" (Costanzo, 2002) but, rather, holds that evolved strategies dealing with information regulation are developmentally based; are initially open-ended but within the boundaries of general constraints, and; become specialized to local social conditions.

In all human cultures, however, revealing certain sensitive information, particularly information involving the personal commission of serious moral transgressions, can be especially damaging to one's genetic interests. Because the punishments for some transgressions can be extreme, including execution, imprisonment, castigation, and isolation, the risks associated with exposing one's involvement in these transgressions are escalated. In such

cases, effectively engaging the information-retention mechanism as a counter-defense against others' information-gathering mechanisms is critical. The lengths an individual goes to in order to guard against social exposure of these misdeeds should notably expand. With so much on the line, individuals might resort to behaviors that, under normal conditions, they would refrain from. As we discuss below, perhaps the most salient example of such behaviors, explicitly governed by information-retention mechanisms, is homicide in the service of protecting confidential information about the self. Once such sensitive details have been revealed, murder may occur if no other tactics are available to prevent the social transmission of this information.

*Information-regulatory mechanisms and homicide*

In Dostoyevsky's (1880/1950) *The Brothers Karamazov*, the Russian monk, Father Zossima, recalls the tale of a contrite and troubled middle-aged man hopelessly attempting to expiate his sin of murdering a young woman many years prior after she refused his marriage offer. The woman's innocent serf had been arrested for her murder, subsequently fell sick in prison, and died shortly thereafter. Plagued by guilt, the man, who "was in a prominent position, respected by all, rich and had a reputation for benevolence" (p. 360) confesses to the priest, but soon comes to regret this public revelation and considers killing Father Zossima: "the thought was unendurable that you were alive knowing everything. . .let me tell you, you were never nearer death" (p. 374).

This fictional scenario allows us to distinguish between the low-level mechanisms driving conspecific killing that are embedded in our primate ancestry, on the one hand, and the evolutionarily novel, high-level mechanisms promoting a new class of adaptive homicidal behaviors characteristic only of humans, on the other hand. The murder of the young woman was sparked by feelings of male proprietariness and jealousy over the woman's engagement to another man. The conscious motives of the murderer's behavior are not central to evolutionary analyses and probability estimates of violence and homicide; no matter what *meaning* the assailant attributes to his actions, his behavior meshes well with the predicted pattern of mate killing and is analogous to similar patterns of aggression and violence in closely related species, such as chimpanzees.

However, the homicidal ideation underlying the second incident, in which the man considers murdering the priest after confessing to the first murder, seems a qualitatively different case in that such thinking demands an appreciation of the mental states of others. The conscious motivation behind such premeditation is important;

should these fantasies have taken shape and the man decided to kill the priest, the murder would be beyond the explanatory power of an evolutionary stance that discounts conscious motives. In human beings, nature capitalizes on the fact that “dead men tell no tales” and motivates affects and behaviors designed to terminate the lives of those who possess damaging information concerning the self. Because these strategies command the resources provided by a theory of mind, which allows for the representation of others’ knowledge and ignorance, the evolutionary scenario provided in this article holds that this is a recent innovation in homicidal behavior and is exclusive to humans. Unlike other forms of murder, which are shared by other primates and which, although they may be entwined with mental representation when manifested in humans (e.g., in the case of infanticide, empathy toward a suffering infant), information-retention homicide *requires* a human theory of mind.

Below, we identify four broad categories of information-retention homicide, each serving to protect the self from the consequences of social exposure in ancestral environments. These include: (1) *first-person exposure, victim-centered homicide*: The perpetrator violates another individual, such as through rape or burglary, and murders the victim to prevent the victim from implicating the perpetrator. The likelihood of homicide occurring in such cases should be a function of the seriousness of the offense, the reaction of the victim during the course of the offense, and the perpetrator’s perceived or actual exposure of his or her identification during the course of the offense. For example, although rape-murder occurs infrequently, the fact that it ever occurs is surprising in that it countermands evolutionary psychologists’ proposed adaptive benefits involved in forced copulation (Thornhill & Palmer, 2000) . However, the adaptiveness of rape-murder becomes clear in light of the current model, which holds that the risk of detection from committing the crime of rape sometimes outweighs the fitness advantages of impregnating a rape victim (see Shackelford, 2002a, 2002b); (2) *second- (or higher) person exposure, witness-centered homicide*: As a consequence of committing some offense, other individual(s) who are not victimized in the offense are exposed to the transgression and to the identity of the perpetrator. Because of this incidental exposure, the perpetrator murders the witness(es) to prevent them from implicating the perpetrator. For example, convenience store robbery-homicides may involve not only the killing of store clerks, but anyone else in the store at the time who happens to see the perpetrator; (3) *defection from mutual criminality, informant-centered homicide*: Two or more individuals participate in a transgression, and one (or more) individual(s) either defects

from the relationship or threatens to confess the transgression. The defector(s) are murdered by the other perpetrator(s) to prevent implication in the transgression. For example, mafia-related killings frequently involve the murder of former business partners, who become “turncoats” or “stoopigeons” or otherwise threaten to testify against current members; (4) *confessional-regret, confessor-centered homicide*: A perpetrator confesses a social transgression to a confidante and then comes to regret divulging this sensitive information. The perpetrator murders the confidante to prevent this individual from transmitting this information. For example, in the context of a marriage, an individual may confess to his spouse a serious social transgression, come to regret this admission when the marriage dissolves, and then decide to murder his or her spouse.

Although not all cases of human murder involve information-retention, many do (although for obvious reasons it may be difficult to get a good estimate on its frequency). In the environment of evolutionary adaptedness, individuals who were able to use these strategies under the threat of social exposure of serious transgressions or undesirable personal traits were more likely to pass on their genes than individuals who were not able to do so. With advances in security systems (e.g., video cameras), forensic science, and communication systems (e.g., telephone, media, the Internet)—each of which can be considered a technological arm of information-gathering mechanisms—the strategy may not be as successful today as it once was. And even in ancestral environments, information-retention homicide demanded retention of information above and beyond the original motive—that is, the murder episode presents the individual with yet additional sensitive information to conceal from others. Still, by *not* committing the murder, social exclusion may be unavoidable. Because perpetrators are *certain* to face grave social consequences if implicated in serious transgressions (such as rape or murder), committing information-retention homicide at least gives their genes a fighting chance of survival.

#### *The natural foundations of confession*

It seems counterintuitive that individuals would engage in confessional behavior when others’ have no knowledge of their transgressions. Because confession guarantees social exposure, it threatens genetic fitness. However, confession may serve to ameliorate the fitness-related blows stemming from what is perceived to be inevitable social exposure. Gold and Weiner (2000) have shown that when confession occurs with remorse signals (such as those accompanying feelings of guilt), others are more likely to infer that the transgression will not occur

again, thus promoting forgiveness and a reduction in retributive behavior. Other authors have speculated that when such signals are in place, people are prone to decide that the party has suffered enough through his guilty feelings, thus making punishment unnecessary. One very testable hypothesis, therefore, is that confession should be most likely to occur when individuals perceive the likelihood of social exposure to be high enough to launch “preemptive strikes” against them.

Also, individuals should be most likely to confess to those who share some stock in their own genetic interests, such as parents and siblings. When confessions are made to those who do not hold such stock, it frequently involves conditions of anonymity (as in the Catholic Church) or confidentiality (as in clinical therapy), both of which are designed to defend against social exposure while satisfying the “urge” to confess. In addition, confession can serve as a signal of commitment to others because it reduces the likelihood of defection from a relationship (e.g., Hong, 1998; Rogers & Holloway, 1993; Shackelford & Buss, 1996; Ting-Toomey, 1991). For example, by confessing, one risks being blackmailed and is therefore more likely to be complicit in a confidante’s schemes (see also Schelling, 1960). A husband’s confession to his wife about a previous affair may be interpreted as a signal of his current emotional commitment, in that if he intended to continue with his extramarital relationship(s) he would not be revealing the indiscretion.

There are important forensic applications to these evolutionary ideas. Although they must be borne out through empirical study, the information-regulatory model enables us to make predictions about confession and information-retention homicide. First, individuals should confess to transgressions under conditions in which social exposure of the offense is perceived as imminent, either subjectively or as a consequence of an overt preponderance of evidence. Therefore, successful interrogation tactics are those that exploit this decision-making heuristic (see Gudjonsson, 2003). For example, authorities may do one of two things to encourage confession when investigating crimes. They may lead the suspect to believe that they possess more information about the criminal activity than they in fact have, or they may give the suspect the impression that they believe what the suspect may have done was socially acceptable (thus dampening the need for information-retention).

We do not envision these processes as being deliberate and consciously strategic. In most cases, confessions should occur as a means to eliminate negative affect, which serves as the proximate mechanism that

led to the ultimate genetic advantages associated with this behavior. Anxiety may be the primary affective response that directly precipitates unbidden confession. Confession has the immediate effect of mitigating such anxiety, and may be the only available behavioral recourse that has this positive effect. Therefore, confession should be the default response under such conditions and should be difficult to inhibit. The foregoing logic presupposes that the subjective degree of precipitating anxiety should be positively correlated with the number of individuals believed to possess knowledge of the offense. This is because social exposure is more likely to occur when more individuals have negative social information about the self. There is some empirical support for this. Bering and Shackelford (2005), for instance, found that federal penitentiary inmates were significantly likely to make unbidden confessions to friends and relatives when they believed that others already knew about their crimes.

From an evolutionary perspective, however, we would also expect that the subjective degree of precipitating anxiety be negatively correlated with the degree to which these individuals are genetically invested in the confessor. All else being equal, one should be more concerned with what their cousin knows about their criminal doings than their brother. Because they are more likely to be associated with genuine remorse, unbidden confessions to the authorities should be associated with greater leniency in punishment than confessions that occur through interrogation tactics. However, it seems probable that individuals first seek others who have some stake in their own genetic interests to share deleterious information before going to the authorities. Because these people have something to lose by the self being punished, genetic kin (or people who share some interest in the individual's genes, e.g., a mate with whom the individual shares common offspring) may have evolved decision-making heuristics that lead them to offer aid in such situations. For example, these people may be more likely than others to come to the individual's physical defense when other ingroup members seek retribution, they may make arrangements for retribution (e.g., offering scarce resources) that serve to negotiate the individual's punishment, or they may help convince ingroup members that the transgression is out of character for the individual and thus unlikely to happen again. Also, with the addition of every genetically invested other, the likelihood of at least one of these people possessing negative social information about punitive ingroup members increases and therefore blackmail may be effectively used to thwart retribution.

If confession occurs and is not soon followed by the anticipated social exposure, or if the precipitating state of anxiety is deemed irrational, then we may expect the individual to feel negative affect toward the person(s) to whom he or she confessed. This negative affect should be driven by the knowledge that this other person is now a carrier of this deleterious information and thus poses a threat to the self's reproductive success. Such negative affect should be heightened whenever the confidante's behavioral response to the confession indicates that he or she will not be complicit in retaining the information from social others, but will likely communicate the information to third parties, thus increasing the probability of social exposure. For example, the confidante may react very negatively toward the confession, verbally impugning the individual for an egregious social violation. Also, a confidante's denial of the seriousness of the offense would be reason to suspect that they will not feel it necessary to retain the deleterious information, and this should also generate negative affect or homicidal ideation toward this person. The degree of negative affect toward the confidante should be equivalent to the social grievousness of the offense. More serious transgressions will be associated with harsher social punishment and therefore more significant damage to the self's genetic fitness. We also predict, however, that any negative affect toward a confidante under such conditions would be negatively correlated with the confidante's degree of genetic relatedness to the confessor. Killing one's genetic kin would dramatically reduce inclusive fitness levels, particularly if the related confidante is high in reproductive value.

#### *Suicide as Failed Information-Regulation*

Under certain conditions, information-retention mechanisms fail to protect sensitive personal information from being publicly disclosed and the individual is faced with social exclusion (e.g., Leary et al., 1995). In such situations, the information-gathering mechanisms of others have won out over the self's retention mechanisms, placing the self's negative attributes on public display. Although some degree of perseverance of the retention mechanisms is apparent (e.g., "hiding one's face in shame"), the damage has been done and there is demand for alternative evolved strategies. The critical evaluations of others—or the anticipation of such evaluations—motivate intense anxiety (Gilbert, 1998), which is scaffolded by negative self-appraisal and shame in response to the "programming" of cultural mores discussed earlier. This negative affect is confounded with the fact that any socio-cognitive meant to change the way others perceive the act (e.g., rationalization, external attributions), are

less effective than information-retention, thus fostering a sense of hopelessness. The confluence of these social, emotional, and cognitive factors has been identified as a very potent and proximal cause of suicide and suicidal ideation (Baumeister, 1990; Kalafat & Lester, 2000; Smith & Hackathorn, 1982).

From an evolutionary perspective, suicide is a challenging phenomenon because it appears at odds with evolved decision-making. Evolutionary processes involve fitting the organism with behaviors enabling it to survive in order to reproduce. However, if the self's survival comes at the expense of inclusive fitness—or, in other words, of genetic kin's ability to pass on genes—then “sacrificing” one's life for ultimate genetic gain may have been adaptive ancestrally (de Catanazaro, 1992). There are numerous cases of “suicide” in other species, particularly among insects, that are compatible with inclusive fitness. For example, it may have been adaptive ancestrally for male Australian redback spiders (*Latrodectus hasselti*) to comply with sexual cannibalism for paternity reasons; males that are cannibalized copulate longer and fertilize more eggs than males that are not cannibalized, and females who have cannibalized males are more likely to reject future suitors than those that have not (Andrade, 1996). The suicidal behavior of pea aphids (*Acyrtosiphon pisum*) parasitized by Braconid wasps (*Aphidius ervi*) is context sensitive, involving apparent calculations of reproductive potential (McAllister, Roitberg, & Weldon, 1990; McAllister & Roitberg, 1987). As reproductive opportunities diminish, pea aphids become more likely to modify their escape behavior when faced with predators such that there is an increased risk of mortality, thus increasing parasite transmission to subsequent hosts in multi-host systems (Holmes & Bethel, 1972). These findings support the “host suicide hypothesis,” which “presumes that the cost of ‘suicide’ may be overridden by the reduced death rate of closely related individuals due to the death of the parasite” (McAllister et al., 1990, p. 167). Another example is the case of bumblebees (*Bombus lucorum*), a species parasitized by conopid flies inserting their larva in the host's abdomen, killing the bumblebee in about 12 days and then pupating until their emergence the following summer (Poulin, 1992). Poulin (1992) has presented evidence that parasitized bumblebees alter their behavior by abandoning the colony, spending all of their time in flower meadows outside the colony where they are susceptible to further parasitoid attacks. By doing so, the affected host leads the flies away from nonparasitized kin, thus increasing inclusive fitness by protecting the colony from infestation. Similar arguments have been posited for suicidal behaviors of birds (e.g., O'Connor, 1978). However, Poulin (1992)



cautions against use of the term “suicide” when describing the self-disadvantageous behaviors of other species, particularly insects. “The adoption of a more dangerous lifestyle by an insect that is bound to die shortly may be adaptive in terms of inclusive fitness, but no more suicidal than, for instance, an ageing animal taking risks to reproduce in the presence of a predator as its inevitable death approaches” (p. 175).

There are no confirmed cases of such behaviors among nonhuman primates. Although there are instances of self-injurious behaviors, usually conflated with stereotypies occurring in abnormal social environments such as laboratories (e.g., Lam, Rupniak, & Iverson, 1991), there is no direct evidence of self-inflicted lethal displays in monkeys and apes. In contrast to these closely related species, suicide accounts for a significant minority of deaths in humans and, in certain age groups, such as adolescence and senescence, it is among the leading causes of death (Pampel & Williamson, 2001; Robbins, Angel, & Kumar, 1981). According to deCatanzaro’s (see 1992; see also Brown et al., 1999) model of self-destruction and preservation, patterns of suicide in humans can be deciphered by applying principles of inclusive fitness. Individuals are most likely to commit suicide when direct reproductive prospects are discouraging and, simultaneously, their survival reduces inclusive fitness by posing a burden to close kin and interfering with their reproductive opportunities. For example, deCatanzaro reports strong correlations between suicidal ideation and individuals’ self-reports of such things as burdensomeness to families, success in heterosexual relations, health problems, homosexuality, number of children, number of friends, loneliness, frequency of sex in the last month, and future financial problems. With increasing age, health problems and burdensomeness replace reproductive potential as primary causal factors in suicidal ideation.

Such evolutionary reasoning casts much-needed light on one of the darkest areas of human behavior. However, although it successfully *interprets* suicidal behavior and ideation within an evolutionary framework, it fails to *explain* the evolutionary processes underlying human suicide adaptations. This is because it does not consider how the proximal cognitive determinants of suicide are deeply embedded in mental representational abilities. The question that we have focused on in this article is whether a comprehensive evolutionary account of human behaviors can be achieved without attending to the unique cognitive processes underlying them, processes that are unshared by other species.

One of the consequences of having the information regulatory mechanisms becoming progressively specialized is that the individual comes to confuse the social jurisdiction of his or her local ecology for a much more abstract deontological system in which culture-based rules of *ought* and *should* become the crucible of moral behavior. Therefore, perceived violation of the rules encourages a view of the self as inherently base, particularly for transgressions or traits for which social repercussions or evaluative appraisals are severe. As Shreve and Kunkel (1991, p. 307) state, “Shame centers the psychological focus on the self, rather than on the behavior.” Or, as Lester (1997, p. 353) writes, the distinction between guilt, which seeks restitution, and shame, which seeks secrecy, can be made as follows: “I can’t believe that I did *that*” (guilt) and “I can’t believe that *I* did that” (shame; see also Gilbert, 1998; Tangney, 2001).

Empirical findings by Lester suggest that shame is a better predictor of suicide than guilt. Similarly, Baumeister (1990, p. 91) describes the causal pathway toward suicidal behavior as including “unfavorable self-attributions, which are broadly meaningful interpretations of self as having stable, undesirable qualities, especially ones that may be predictive of additional difficulties in the future.” Other evidence suggests that particularly heightened displays of self-awareness precipitate suicidal behavior. Baumeister (1990) has even referred to suicide as “escape from self.” For example, suicide notes are notorious for containing first-person singular pronouns, and when suicide note writers do mention significant others, they speak of them as being cut off or distant (see Baumeister, 1990).

What therefore distinguishes suicide in humans from the self-disadvantageous behaviors of other species is that only the former occurs in the context of social relations and is not caused by threats of interspecies predation. Rather, it is the threat of intraspecies negative social appraisal and the self’s agreement with these appraisals that lead to the majority of suicidal behavior and ideation in humans. As in the case of information-retention homicide, however, the role of mental representation is vital to evolutionary models of suicide. The chronic and acute anguish resulting from these socio-cognitive processes might have been specially designed to promote suicide, a behavior that preserved inclusive fitness in the environment of evolutionary adaptedness.

*Concluding remarks*

Part of the reason that evolutionary psychology has been reluctant to address the role of higher-order cognitive mechanisms in human evolution may be that it is averse to the concept of *conscious motive*, which is often conflated to mean that people deliberately engage in particular behaviors because they are genetically adaptive. Evolutionary psychology is right to be dismissive when higher-order cognition is discussed in this causal sense. But the distinction that we have made in this article is a subtle one. For example, in information-retention homicide, an individual knows that he or she wants to kill another person because that person possesses knowledge that is harmful to the self; indeed these higher-order cognitive abilities can *cause* the adaptive behavior. This is not the same as saying that the individual knows that she wants to kill someone because that person's death will salvage her genetic fitness. Nevertheless, ours is a qualitatively different model than those arguing for the relative causal unimportance of higher-order cognition in human social evolution.

We should end by saying that we are not claiming that all ancestrally-adaptive human behaviors stem from the coevolution of theory of mind and language. The majority of human adaptations may have nothing at all to do with mental state representation. But there is compelling enough reason, we think, to suspect that everything that makes us uniquely human...does.

References

- Ahmad, Y. & Smith, P. K. (1994). Bullying in schools and the issue of sex differences. In J. Archer (Ed.), *Male violence* (pp. 70-83). London: Routledge.
- Atzwanger, K. (1995, June). *Biological aspects of driving behavior*. Paper presented to the Annual Conference of the Human Behavior and Evolution Society, Santa Barbara, CA.
- Baumeister, R. F. (1990). Suicide as escape from self. *Psychological Review*, 97 (1): 90-113.
- Bering, J. M., & Shackelford, T. K. (2005). The evolutionary psychology of false confession. *American Psychologist*, 60(9): 1037-1038.
- Bering, J. M., & Shackelford, T. K. (2004). The causal role of consciousness: A conceptual addendum to human evolutionary psychology. *Review of General Psychology*, 8: 227-248.
- Bjorklund, D. F. (1997). The role of immaturity in human development. *Psychological Bulletin*, 122 (2): 153-169.
- Bjorklund, D. F. & Pellegrini, A. D. (2002). *The origins of human nature: Evolutionary developmental psychology*. Washington, DC: American Psychological Association.
- Brown, R. M., Dahlen, E., Mils, C., Rick, J., & Biblarz, A. (1999). Evaluation of an evolutionary model of self-preservation and self-destruction. *Suicide & Life-Threatening Behavior*, 29 (1): 58-71.
- Buss, D. M. (1999). *Evolutionary psychology: The new science of the mind*. Boston: Allyn and Bacon.
- Buss, D. M., & Dedden, L. A. (1990). Derogation of competitors. *Journal of Social and Personal Relationships*, 7 (3): 395-422.
- Buss, D. M., Haselton, M. G., Shackelford, T. K., Bleske, A. L. & Wakefield, J. C. (1998). Adaptations, exaptations, and spandrels. *American Psychologist*, 53 (5): 533-548.
- Campbell, A. (2002). *A mind of her own: The evolutionary psychology of women*. London: Oxford University Press.
- Choi, Y-S. (1996). Family views about social stigma of mental illness patients. *Korean Journal of Clinical Psychology*, 15 (1): 69-89.

- Clements, P. T. & Burgess, A. W. (2002). Children's responses to family member homicide. *Family & Community Health*, 25 (1): 32-32.
- Costanzo, P. (2002). Social exchange and the developing moral syntax of moral orientation. In B. Laursen and W. G. Graziano (Eds.), *Social exchange in development. New directions for child and adolescent development* (pp. 41-52). San Francisco: Jossey-Bass/Pfeiffer.
- Cutting, A. L. & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development*, 70 (4): 853-865.
- Crozier, W. R. & Skliopidou, E. (2002). Adult recollections of name-calling at school. *Educational Psychology*, 22 (1): 113-124.
- Daly, M. & Wilson, M. (1988). *Homicide*. Hawthorne, NY: Aldine de Gruyter.
- Dawkins, R. (1986). *The blind watchmaker*. New York: Norton.
- deCatanzaro, D. (1992). Prediction of self-preservation failures on the basis of quantitative evolutionary biology. In R. W. Maris, A. L. Berman, J. T. Maltzberger, & R. I. Yufit (Eds.), *Assessment and prediction of suicide* (pp. 607-621). New York: Guilford Press.
- Dennett, D. C. (1995). *Darwin's dangerous idea*. New York: Simon & Schuster.
- Dostoyevsky, F. (1950). *The brothers Karamazov*. Trans by C. Garnett. New York: Modern Library.  
(Original work published 1880)
- Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, 16 (4): 681-735.
- Dunn, R. G. & Morrish-Vidners, D. (1988). The psychological and social experience of suicide survivors. *Omega: Journal of Death & Dying*, 18 (3): 175-215.
- Flavell, J. H. (1999). Cognitive development: Children's knowledge about the mind. *Annual Review of Psychology*, 50: 21-45.
- Gabel, S. (1992). Children of incarcerated and criminal parents: Adjustment, behavior, and prognosis. *Bulletin of the American Academy of Psychiatry & the Law*, 20 (1): 33-45.

- Gangestad, S. W. & Cousins, A. J. (2001). Adaptive design, female mate preferences, and shifts across the menstrual cycle. *Annual Review of Sex Research*, 12, 145-185.
- Gigerenzer, G. (2001). The adaptive toolbox: Toward a Darwinian rationality. In J. A. French, A. C. Kamil, and D. W. Leger. (Eds.), *Evolutionary psychology and motivation*. Vol. 47 of the *Nebraska Symposium on Motivation* (pp. 113-143). Lincoln, NE: University of Nebraska Press.
- Gilbert, P. (1998). What is shame? Some core issues and controversies. In P. Gilbert and B. Andrews (Eds.), *Shame: Interpersonal behavior, psychopathology, and culture*. Series in affective science (pp. 3-38). New York: Oxford University Press.
- Gold, G. J. and Weiner, B. (2000). Remorse, confession, group identity, and expectancies about repeating a transgression. *Basic and Applied Social Psychology*, 22 (4): 291-300.
- Gladue, B. A. (1991). Qualitative and quantitative sex differences in self-reported aggressive behavioral characteristics. *Psychological Reports*, 68 (2): 675-684.
- Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions: A handbook*. Wiley series in the psychology of crime, policing and law. New York: John Wiley & Sons.
- Hammersmith, S. K. (1987). A sociological approach to counseling homosexual clients and their families. *Journal of Homosexuality*, 14 (1-2): 173-190.
- Herek, G. M. (1998). *Stigma and sexual orientation: Understanding prejudice against lesbians, gay men, and bisexuals*. Newbury Park, CA: Sage.
- Holmes, J. C. & Bethel, W. M. (1972). Modification of intermediate host behavior by parasites. In E. U. Canning and C. A. Wrights (Eds.), *Behavioural aspects of parasite transmission* (pp. 123-149). London: Academic Press.
- Hong, D. S. (1998). Variables related to formation and development of dating relationships. *Korean Journal of Social & Personality Psychology*, 12 (2): 199-266.
- Humphreys, J. C. (2001). Growing up in a violent home: The lived experience of daughters of battered women. *Journal of Family Nursing*, 7 (3): 244-260.
- Kalafat, J. & Lester, D. (2000). Shame and suicide: A case study. *Death Studies*, 24 (2): 157-162.

- Lam, K., Rupniak, N. M. J., & Iverson, S. D. (1991). Use of a grooming and foraging substrate to reduce cage stereotypies in macaques. *Journal of Medical Primatology*, 20: 104-109.
- Leary, M. R., Tambor, E. S., Terdal, S. K., & Downs, D. L. (1995). Self-esteem as an interpersonal monitor: The sociometer hypothesis. *Journal of Personality and Social Psychology*, 68 (3): 518-530.
- Lester, D. (1997). The role of shame in suicide. *Suicide & Life Threatening Behavior*, 27 (4): 352-361.
- Lillard, A. S. (1997). Other folks' theories of mind and behavior. *Psychological Science*, 8 (4): 268-274.
- Lillard, A. S. (1998). Ethnopsychologies: Cultural variations in theory of mind. *Psychological Bulletin*, 123 (1): 3-32.
- May, H. (2000). "Murderers' relatives": Managing stigma, negotiating identity. *Journal of Contemporary Ethnography*, 29 (2): 198-221.
- McAllister, M. K. & Roitberg, B. D. (1987). Adaptive suicidal behaviour in pea aphids. *Nature*, 328 (6133): 797-799.
- McAllister, M. K., Roitberg, B. D., & Weldon, K. L. (1990). Adaptive suicide in pea aphids: Decisions are cost sensitive. *Animal Behaviour*, 40 (1): 167-175.
- McAndrew, F. T. & Milenkovic, M. A. (2002). Of tabloids and family secrets: The evolutionary psychology of gossip. *Journal of Applied Social Psychology*, 32 (5): 1064-1082.
- Miall, C. E. (1987). The stigma of adoptive parent status: Perceptions of community attitudes toward the adoption and the experience of informal social sanctioning. *Family Relations: Journal of Applied Family & Child Studies*, 36 (1): 34-39.
- Murphy, D. A., Roberts, K. J., & Hoffman, D. (2002). Stigma and ostracism associated with HIV/AIDS: Children carrying the secret of their mothers' HIV+ serostatus. *Journal of Child & Family Studies*, 11 (2): 191-202.
- O'Connor, R. J. (1978). Brood reduction in birds: Selection for fratricide, infanticide and suicide. *Animal Behaviour*, 26 (1): 79-96.
- Pampel, F. C. & Williamson, J. B. (2001). Age patterns of suicide and homicide mortality rates in high-income nations. *Social Forces*, 80 (1): 251-282.
- Perner, J., Ruffman, T., & Leekam, S. R. (1994). Theory of mind is contagious: You catch it from your sibs. *Child*

*Development*, 65 (4): 1228-1238.

Peterson, C. C. (2001). Influence of siblings' perspectives on theory of mind. *Cognitive Development*, 15 (4): 435-455.

Peterson, C. C. & Siegal, M. (2002). Mindreading and moral awareness in popular and rejected preschoolers. *British Journal of Developmental Psychology*, 20 (2): 205-224.

Poulin, R. (1992). Altered behaviour in parasitized bumblebees: Parasite manipulation or adaptive suicide? *Animal Behaviour*, 44 (1): 174-176.

Povinelli, D. J., Bering, J. M., & Giambrone, S. (2000). Toward a science of other minds: Escaping the argument by analogy. *Cognitive Science*, 24 (3): 509-541.

Robbins, S. L., Angel, M., & Kumar, V. (1981). *Basic pathology* (3<sup>rd</sup> ed). Philadelphia: Saunders.

Rogers, J. C. & Holloway, R. L. (1993). Professional intimacy: Somewhere between collegiality and personal intimacy? *Family Systems Medicine*, 11 (3): 263-270.

Ruffman, T., Perner, J., & Parkin, L. (1999). How parenting style affects false belief understanding. *Social Development*, 8 (3): 395-411.

Schelling, T. C. (1960). *Strategy of conflict*. Cambridge, MA: Harvard University Press.

Schwanenflugel, P. J., Martin, M., & Takahashi, T. (1999). The organization of verbs of knowing: Evidence for cultural commonality and variation in theory of mind. *Memory & Cognition*, 27 (5): 813-825.

Shackelford, T. K. (2002a). Are young women the special targets of rape-murder? *Aggressive Behavior*, 28 (3): 224-232.

Shackelford, T. K. (2002b). Risk of multiple-offender rape-murder varies with female age. *Journal of Criminal Justice*, 30 (2): 135-141.

Shackelford, T. K., & Buss, D. M. (1996). Betrayal in mateships, friendships, and coalitions. *Personality and Social Psychology Bulletin*, 22 (11): 1151-1164.

Shreve, B. W. & Kunkel, M. A. (1991). Self-psychology, shame, and adolescent suicide: Theoretical and practical considerations. *Journal of Counseling & Development*, 69 (4): 305-311.

Smith, D. H. & Hackathorn, L. (1982). Some social and psychological factors related to suicide in primitive



- societies: A cross-cultural comparative study. *Suicide & Life-Threatening Behavior*, 12 (4): 195-211.
- Tangney, J. P. (2001). Self-conscious emotions: The self as a moral guide. In A. Tesser, D. A. Stapel, and J. V. Wood (Eds.), *Self and motivation: Emerging psychological perspectives* (pp. 97-117). Washington, D. C.: American Psychological Association.
- Ting-Toomey, S. (1991). Intimacy expressions in three cultures: France, Japan, and the United States. *International Journal of Intercultural Relations*, 15 (1): 29-46.
- Tooby, J. & Cosmides, L. (1992). The psychological foundations of culture. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 19-136). New York: Oxford University Press.
- Veltman, A., Cameron, J. I., & Stewart, D. E. (2002). The experience of providing care to relatives with chronic mental illness. *Journal of Nervous & Mental Disease*, 190 (2): 108-114.
- Wellman, H. M. (1998). Culture, variation, and levels of analysis in folk psychologies: Comment on Lillard (1998). *Psychological Bulletin*, 123 (1): 33-36.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72 (3): 655-684.
- Williams, G. C. (1992). *Natural selection*. New York: Oxford University Press.