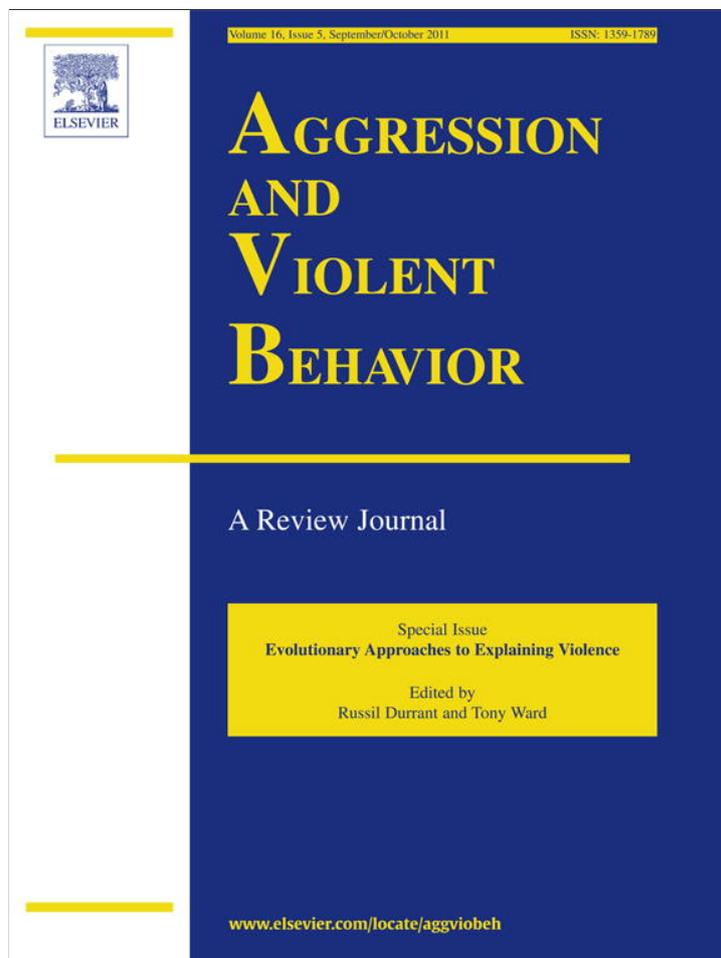


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Aggression and Violent Behavior



The recalibrational theory and violent anger

Aaron N. Sell*

Griffith University, Australia

ARTICLE INFO

Article history:

Received 15 February 2011
 Received in revised form 7 April 2011
 Accepted 11 April 2011
 Available online 20 April 2011

Keywords:

Recalibrational theory
 Anger
 Evolutionary psychology
 Aggression

ABSTRACT

Anger is responsible for a large share of human aggressive acts, and yet most incidents of anger do not result in any aggressive behavior at all. To understand when and why an angry individual becomes violent we must understand what anger is and what function it solves. The recalibrational theory of anger is a computational-evolutionary model that maintains that the function of anger is to recalibrate individuals who place insufficient weight on the welfare of the angry individual when making decisions, i.e. anger raises the target's welfare tradeoff ratio (WTR). Features of anger that can be explained from this perspective include the major triggers of anger, the content of anger-based arguments, individual differences in anger thresholds, and the selective deployment of negotiative tactics such as the withdrawal of benefits and the infliction of costs. When certain criteria are met, cost infliction will take the form of violent aggression and can be understood as a bargaining tactic designed to recalibrate the target's WTR.

Crown Copyright © 2011 Published by Elsevier Ltd. All rights reserved.

Contents

1. Conflicts of interest and welfare tradeoff ratios (WTRs)	382
2. The recalibrational theory of anger	382
2.1. What are the triggers of anger?	383
2.1.1. Cost imposition as a trigger of anger	383
2.1.2. Insults	383
2.2. Stage 1 anger	384
2.3. Stage 2 anger and WTR negotiation	384
2.3.1. Negotiating with benefit withdrawal	385
2.3.2. Negotiating with cost infliction	385
2.3.3. Choosing among negotiating tactics	386
3. A psychologist's guide to eliciting violent anger	387
4. Conclusions	388
References	389

Living tissue is made of dead matter (mostly hydrogen, oxygen and carbon), but is “alive” to the extent that it is functionally designed to bring about improbable outcomes. Aggression is, analyzed unemotionally and with the full breadth of the animal kingdom in mind, the disordering of another animal in ways that prevent its long term functioning. From the standpoint of natural selection, disordering other organisms would have been an attractive option in a wide range of circumstances (Archer, 1988; Duntley & Buss, 2004). But the targets of these attempts are, of course, designed to resist threats to their

functional integrity often by disordering the “attacker.” This makes aggression a dangerous tactic.

Because of this danger, the mechanisms that natural selection has designed to deploy aggression use that option only in limited circumstances and often only after less dangerous tactics have failed. To understand when aggression is deployed, we require a careful examination of the full range of conflict adaptations not just the act of aggression itself. To only study *aggression* per se would be rather like trying to study courtship by looking at the mechanics of sexual reproduction. Much of the negotiating, assessment, and decision making has already been done by that point. Likewise, to understand why an act of aggression took place, you must study the nature of conflict mechanisms that are online and making computations and

* School of Criminology and Criminal Justice, Griffith University, Mt. Gravatt Campus, Australia.

E-mail address: a.sell@griffith.edu.au.

informing motivational systems from the moment two individuals meet. These mechanisms are ultimately what give a satisfying explanation for acts of aggression.

There are likely many mental systems which deploy aggression, and each will respond to their own input conditions and moderating computational variables. For example, defensive aggression against would-be predators likely comes from a different neurological source than hate-filled homicidal aggression (Moyer, 1968). These different systems will be the result of the differing selection pressures put on the mind by the structure of the natural world, and we can therefore look at the nature of the adaptive problems our ancestors faced and posit a taxonomy of mental systems accordingly. This taxonomy will no doubt be refined as we map the computational structure of each of the systems, e.g. does the aggression directed at disobedient offspring result from the same mental mechanisms as aggression directed at disrespectful peers? (see Bugental, Blue, & Cruzcosa, 1989). Herein I will explore one mental system that I believe is responsible for a great deal of human aggression, specifically anger and the recalibrational aggression that it can deploy. Before the theory can be elaborated, however, we need to explore the selection pressures that are responsible for the evolution of this system.

1. Conflicts of interest and welfare tradeoff ratios (WTRs)

In social species, decisions by one individual will frequently affect the welfare of others. As a result, natural selection can be expected to have designed mental mechanisms that determine how much weight an individual should put on the welfare of another compared to themselves when making decisions that affect them both. The mental representation of this relative weighting has been called a *welfare tradeoff ratio* (WTR; Tooby, Cosmides, Sell, Liberman, & Sznycer, 2008) because it sets the ratio of values below which an individual will tradeoff another's welfare for their own benefit. One's WTR toward another is revealed by conflicts of interest wherein a person must choose between two mutually incompatible outcomes with different consequences for each individual. This can be expressed mathematically as: $V_x > WTR_{xy} * V_y$, which is the decision rule used by agent X when making a decision about whether to take something of differing value (V) to both X and Y. For example, if X held a WTR toward her graduate student (Y) of 1/2, then X would use the rule $V_x > 1/2 V_y$, indicating that X would take the self-beneficial action unless the value of the resource to Y was more than twice the value to herself. In other words, X would value Y's welfare as half of her own and would impose a cost of 6 on him to receive a benefit of 4 but not a benefit of 2. Welfare tradeoff ratios are person-specific and underlie many human concepts such as respect, deference, regard and love.

In general, natural selection can be expected to have designed animals to calibrate their welfare tradeoff ratios to maximize their own welfare.¹ If biting a conspecific's offspring to death enhances an individual animal's reproductive prospects, natural selection will reward that (Sherman, 1981). However, if a conspecific is capable of retaliating, or if imposing costs on a conspecific leads them to be unwilling or unable to continue cooperating, then natural selection will reward setting a higher WTR toward that individual. In particular, an individual would be selected to raise their WTR toward another when doing so would be less costly than maintaining the lower WTR and absorbing the costs of conflict and the denial of cooperative benefits. In short, animals negotiate their respective WTRs by leveraging 1) their ability to impose costs and 2) their ability to confer benefits.

One way to impose costs on others is to deploy aggression against them. Indeed, species that engage in frequent conspecific aggression

calibrate their behavior to information about relative fighting ability. Many such species assess fighting ability from visual, auditory or olfactory cues, store these assessments over time, use transitive reasoning to determine relative fighting ability, calibrate their assessments based on victory or lose in a fight, and engage in aggressive rituals to test relative formidability, sometimes even in the absence of a current conflict of interest (Huntingford & Turner, 1987; Sell, 2012). Many species, particularly social primates, also assess alliances and cooperative value and show deference (i.e. set higher WTRs) toward individuals with more allies or with the ability to deny benefits (Smuts, Cheney, Seyfarth, Wrangham, & Struhsaker, 1987). This mutual calibration allows animals to resolve conflicts of interests in the absence of repeated combat and results in the formation of stable patterns of interactions sometimes called dominance hierarchies.

In short, humans and other animals assess traits relevant to the ability to impose costs (e.g. fighting ability, coalitional size) and bestow benefits (e.g. mate value, resource gathering potential) and use these estimates to calibrate their welfare tradeoff ratios such that more weight is placed on the welfare of individuals who can efficiently impose costs or bestow benefits on others. During human evolutionary history, others' WTRs toward you would have been a powerful predictor of your ability to receive food, acquire and retain mates, receive aid while sick, protect ones offspring from injury and predation, and influence others. Given the substantial impact of other individuals' WTRs on one's fitness, natural selection can be expected to have designed a mechanism to interface with WTR-setting machinery in others and recalibrate them so as to raise their WTR. According to the recalibrational theory, anger is this mechanism.

2. The recalibrational theory of anger

A computationally detailed theory of anger has been missing from emotional psychology because anger is a complex emotion. It has numerous triggers including insults, cost imposition, inattention, anger from another, insufficient reciprocity, insufficient praise, another's ignorance of your achievements and so on. In response, anger coordinates physiological reactions, facial expressions, vocal responses, visual attention, body postures, and a host of behavioral strategies including social exclusion, insults, argumentation and aggression. A successful theory of anger will have to account for all of these features in addition to predicting and explaining new features of anger that are currently undiscovered.

The recalibrational theory of anger holds that a large and well-bound subset of what is called anger is the result of an adaptation designed by natural selection that functions primarily to recalibrate another individual who has demonstrated a lower than acceptable welfare tradeoff ratio toward the angry person so that, in the present and future, the target of anger will weigh the angry individual's interests more highly, i.e. raise their WTR.² This model makes specific predictions, some of which have been tested already, about the triggers of anger, the kinds of computations run by the anger system, the source of individual differences in anger, the relationship between entitlement and anger, and how anger and anger-based aggression can be neutralized. It is broadly consistent with theories that view anger as a mechanism of impression management (Felson, 1982), self-protection (Lazarus, 1991), or commitment devices that deploy retaliatory aggression (Frank, 1988; Hirshleifer, 1987). However, by

¹ This is complicated by kinship and "deep engagement" relationships in which individual's genetic interests overlap to the point where one individual has a genuine interest in the survival and well being of another (e.g. individuals with shared offspring).

² The word "anger" like the word "love" refers to many phenomena that are not likely to have the same neurocomputational explanations. The anger described in this paper is interpersonal anger and does not refer to frustration at inanimate objects (though computers do mimic human interactions in ways that might elicit it), anger at oneself (i.e. regret) or hatred.

taking an explicitly evolutionary-functional approach, the recalibrational theory is capable of making precise a priori predictions about many aspects of anger and how those aspects function differently in cooperative versus antagonistic relationships.

2.1. What are the triggers of anger?

According to the recalibrational theory, anger is activated by information that another individual holds a welfare tradeoff ratio toward the angry individual that is too low. This information can come in many formats such as insulting beliefs, disregard of an individual's opinions or beliefs, cost imposition, and so on. Each trigger of anger will have to be unpacked individually to reveal its cognitive structure, and often will depend on cognitive mechanisms that serve different functions, e.g. anger can be triggered when someone does not remember information about the angry individual; to understand this trigger of anger we will require a computational account of person-specific memory and why, presumably, it uses WTRs to regulate memory space. Two examples from the most common triggers of anger will be outlined here.

2.1.1. Cost imposition as a trigger of anger

Welfare tradeoff ratios are used primarily when making decisions about whether to tradeoff on someone else's welfare to benefit oneself. Naturally then, such interactions are one way to assess the magnitude of another's WTR. If X forgoes a large benefit in order to avoid imposing a trivial cost on Y, then X has demonstrated a high WTR toward Y. On the other hand, if X extracts a small benefit by imposing a large cost on Y, then X has demonstrated a low WTR toward Y. So while the most common trigger of anger is the imposition of a cost on the angry person, not all cost impositions are sufficient to trigger anger. Imposing a small cost for a large benefit is still consistent with a high WTR. For example, if your fawningly obedient graduate student has a WTR of 3:1 in your favor, he would still impose a cost on you if he benefited 4 times as much as you would be harmed: $V(\text{grad}) > 3 V(\text{you})$. By definition, a lower welfare tradeoff ratio is revealed by a cost imposition as: a) the cost imposed *increases*, and b) the benefit reaped by the imposer *decreases*.

Indeed, past research shows conclusively that the probability of triggering the anger response is positively related to the magnitude of the cost. Empirical studies have confirmed this across a host of different types of costs, including electric shocks (Taylor, 1967) severity of insults (Harmon-Jones & Sigelman, 2001), crimes (based on sociological survey data on attitudes about crime, e.g. White, 1975), and monetary payoffs (Fehr & Gaechter, 2000). More recent research also shows that costs imposed for particularly large benefits are less likely to trigger anger even when the costs themselves are substantial. This same pattern has been replicated across numerous cultures (Sell et al., forthcoming). The data suggest it is the cue of a low WTR that is triggering anger, not the cost itself.

This view is bulwarked by research on the role intentionality plays in mediating the relationship between cost-imposition and anger. In order for a cost imposition to indicate a low WTR, three elements must be known to the person who imposed the cost: i) the magnitude of the cost imposed, ii) the magnitude of the benefit the other gained by imposing that cost, and iii) the identity of the individual on whom that cost was imposed. If the individual who imposed the cost misperceived the magnitude of the benefit as higher than it was or the cost as lower than it was, then it is possible that their WTR is high enough to not trigger anger. Additionally, because welfare tradeoff ratios are person-specific, the imposition of a large cost for a low benefit will clearly indicate a low WTR toward an individual only if it was known that the cost would be imposed on that specific individual. A long research tradition shows that unintentional costs are much less likely to trigger anger (e.g. Ohbuchi and Kambara, 1985), though the

specific effect of each of the three predicted elements has not yet been tested.

The view that cost-impositions are clues to WTRs rather than noxious stimuli in themselves offers an explanation for the rather puzzling phenomena that homicides and assaults are frequently the result of conflicts over trivial resources (Daly & Wilson, 1988). The absolute cost of the interaction does not have to be large for it to indicate the relative weight a person puts on the angry individual's welfare compared to their own. If someone feels entitled to push you out of the line at the grocery store because they don't feel like waiting, that same person will do much worse to you when the stakes are higher. None of this need be conscious in the mind of the angry individual, but that situation – that individuals with low WTRs will impose large costs on you over time – selected for an anger system that responds to cues, even cues resulting from trivial interactions, that another individual devalues your interests highly.

2.1.2. Insults

Because WTRs are set in response to assessments of various traits relating to how efficiently an individual can impose costs or confer benefits, an indication that X believes Y to be deficient on such a trait also means that X is setting his WTR toward Y at a lower level than he would with accurate information. For example, because a man's physical formidability is used when setting WTRs (Sell, Tooby, & Cosmides, 2009), if a man discovers that his neighbor believes him to be physically weak then it is likely that his neighbor has a lower WTR toward the man than he would if he knew he was not weak. This should hold for any negotiable feature that is used to set welfare tradeoff ratios: physical strength, sexual fidelity, attractiveness, wealth, intelligence, skills and talents, reliability as a cooperater, and any other factor that relates to one's ability to negotiate for better treatment. Revealing one's low estimation of another's WTR-relevant trait can be called an "insult." Insults have been used as reliable triggers of anger in aggression research for forty years (Geen, 1998). Indeed, research on the most ecologically valid examples of aggression has found that personal "insults" almost always precede homicides (Berg & Fox, 1947; Daly & Wilson, 1988; Luckenbill, 1977) and assaults (Felson, 1982).

The word "insult" is often used for explicit verbal degradations, but by the logic of the model anger should be triggered by any indication that another represents the angry person as being deficient on some WTR-relevant trait. Presumably the probability of triggering anger is proportional to the gap between the estimates of the angry individual and the target (e.g. receiving a rating of 3 on a 9-point scale of respect is more anger provoking than being rated as a 7, Harmon-Jones & Sigelman, 2001) and the degree to which the trait in question affects the WTR's magnitude, (e.g. believing a man is afraid of physical violence is more insulting to him than it would be to a woman because men are valued for their martial courage [Harris, 1993]).

A prediction of this model is that insults will cluster largely around the two primary negotiative tools that are used to set WTRs: the ability to impose costs (e.g. being weak, cowardly, socially impotent) and the ability to confer benefits (e.g. being unreliable, a cheat, unable to gather resources, having low mate value, being incompetent, being sickly, being a burden on others). Another prediction from this model is that what is insulting to one person will not be insulting to another, depending upon how WTRs are set for a person in their situation. There is abundant evidence, for example, that anger is triggered by different insults in men and women depending upon how they are valued. Adult men, for example, compete and bargain for higher WTRs, in part, by using physical formidability (Sell et al., 2009; von Rueden, Gurven, & Kaplan, 2008) and so they find it insulting to be thought of as weak or cowardly. Adult women are valued in part for their physical attractiveness and sexual fidelity and thus find negative beliefs about their attractiveness or sexual history highly insulting (Harris, 1993; Preston & Kimberley, 1987).

Note that when anger is triggered in this way, the issue of intentionality disappears. If a friend were to indicate that he thought your last paper was rubbish, your anger would not be easily dispelled by statements such as, “but I didn't know your paper would be terrible” or “I never intended to think your paper was rubbish.” Those statements are all but incoherent. Distinctions such as these indicate the importance of studying anger as a functional system with a somewhat modular design in which crucial variables for one trigger of anger become wholly unimportant for another. By seeing cost-impositions and insulting beliefs as two separate gauges on another's WTR, however, the recalibrational theory of anger explains these distinctions neatly.

Because insulting beliefs are a rapid and cheap way of demonstrating a low WTR, they can be used when individuals want to challenge someone, i.e. to indicate that they do not value that individual and welcome their negotiations. In such cases derogatory statements are made explicitly for the purpose of triggering the anger mechanism (e.g. “fuck you” or “your mother's a whore”). These are not meant to be taken literally; indeed few make sense when thought of that way. Instead they are challenges that work by indicating to the other that you intend to not value them highly. This can be done unprovoked (e.g. verbal bullying, see Swain, 1998) or as the first bid in an angry negotiation (e.g. someone implies you are not a hard worker and you swear at them, see case 34, Luckenbill, 1977). Similar challenges can be issued using cost inflictions with high costs for trivial or no benefits, e.g. spitting on someone.

2.2. Stage 1 anger

Anger is an expensive emotion, both metabolically (Ekman, Levenson, & Friesen, 1983) and measured in terms of the loss of cooperation that can result from its frequent deployment. This is particularly true when speaking of the primary tools that anger has available to it, specifically cost-infliction and the withdrawal of benefits. Because of this, and because intuiting another's WTR is always an imperfect estimate, the first response of the anger system is usually to seek information from the target to refine their WTR estimate and give the target of anger a chance to explain why their actions do not indicate a low WTR. I call this *stage 1 anger*. This information exchange is what characterizes anger-based arguments, the most common behavioral response to an incident of anger (Averill, 1982).

The kinds of probing questions asked, and the kinds of arguments offered during stage 1 anger will depend on the triggering incident because each trigger of anger responds to different variables. When the anger system is triggered by cost impositions, for example, it seeks out information that allows for a more accurate estimation of the target's WTR. Such information is, as predictable from the model itself, the target's estimates of the variables in the WTR equation, i.e., the target's knowledge of the magnitude of the cost to the angry person (e.g. “do you know how much that hurt me?”) and the magnitude of the benefit (e.g. “why did you do that?”). When evidence exists that the target perceives a cost or benefit differently than the angry individual, the anger system should be designed to modify the target's perception of that variable in ways that prevent this particular transaction and any future subset of transactions involving the same misconstrued variable. Such “debates” about the relative value of a contested resource are not unique to humans; e.g. during a conflict over a food source a bald eagle will lift its head and display its non-distended crop to demonstrate that it has not recently fed and thus values the food item highly (Hansen, 1986).

Because stage 1 anger is designed to modify estimates of variables that relate to WTR expression (e.g. magnitude of costs, intentionality) rather than welfare tradeoff ratios per se, the negotiative tactics are quite distinct. For example, take a young married couple (Mrs. X and Mr. Y) who are arguing over whether to attend the man's uncle's

funeral which he wants (Vy), or go out on their anniversary dinner which she wants (Vx). They would be expected to argue over Vx, i.e. the cost of not having the dinner. She would argue that Vx is high, e.g. “we already have reservations” “we don't see each other much anymore”, “I've been looking forward to this for months” and he would argue that Vx is low, e.g. “we can go next week instead”. Similar arguments would characterize Vy, i.e. the funeral; she would argue Vy is not large, e.g. “you'd have to fly across the country; you hate flying” and he would argue Vy is large, e.g. “my family will all be there.”

Much of the data gathered by Averill's (1982) extensive survey of anger is consistent with this interpretation. For example, anger-based arguments resulted in reinterpretations of relevant variables for 62% of the instances of anger. Most commonly, the angry person reinterpreted the motives or guilt of the target of anger (e.g. decided the target received a high benefit or had acted unintentionally). The second most common reinterpretation by the angry individual was that he or she “decided the incident was less important than originally thought” (e.g. low cost to themselves). The targets of anger also responded in ways consistent with the recalibrational theory. According to Averill (1982, pg. 223), the most common long-term response on behalf of the target of anger was that they, “realized their own faults.” Fully 76% of the targets of anger agreed that they realized their own faults either “somewhat” or “very much.” In other words, anger has a high success rate even when not deploying aggression (less than 5% of Averill's subjects reported any aggressive act).

When these early tactics do fail, however, the angry person must decide between tolerating the low WTR or switching tactics and escalating negotiations. This decision is likely made by weighing the future costs of the target's low WTR against the costs of recalibrating him. The likelihood of future interaction will no doubt be one important variable determining those costs, as will factors about the kind of reputational damage done by ignoring affronts.

What is absent from stage 1 anger is negotiative tactics that are relevant for setting WTRs, such as one's ability or willingness to impose costs or remove benefits. Notice, for the example above, how the argument would be qualitatively changed if Mr. Y were to say, “I could beat the hell out of you and no one would do anything” or if Mrs. X were to say, “Go then, I could always get your boss to take me out for a romantic dinner.” Such statements are not relevant to a debate about the relative value of some course of action to the individuals involved, but are rather bargaining tactics meant to indicate one's ability to impose costs or remove benefits from the other. In other words, those arguments are only deployed by the anger system when it is functioning to modify the target's WTR rather than attempting to change a course of action, i.e. during stage 2 anger.

2.3. Stage 2 anger and WTR negotiation

In response to clear evidence that a target has a lower than acceptable welfare tradeoff ratio, the anger system will deploy tactics that stimulate the input conditions of the target's WTR-setting machinery, i.e. anger attempts to recalibrate the target's WTR upward. I call this *stage 2 anger* – the system is now engaging in a more high stakes negotiation over long-term changes in the behavioral patterns of the target of anger rather than negotiating about a particular conflict of interest. To recalibrate a WTR, the anger system must feed information to the WTR-setting machinery in a format that is computationally intelligible to that system, i.e. because WTRs are set only by certain classes of variables, those specific variables must be demonstrated to the target in order to recalibrate him. Because WTRs are largely set by clusters of variables that fall into two categories – the ability and willingness to impose costs, and the ability and willingness to withdraw benefits – these are the tactics with which stage 2 anger can negotiate.

2.3.1. Negotiating with benefit withdrawal

The mind should be designed to track cues of the ability to confer benefits and respond automatically by awarding more consideration to the welfare of individuals from whom one has received benefits (Tooby & Cosmides, 1996). Furthermore, individuals who have the potential to deliver great benefits to others should, and do, spontaneously expect better treatment (i.e. demand a higher WTR) and set lower thresholds for anger (Sell et al., 2009).

Because welfare tradeoff ratios are set in part by the relative ability of individuals to confer benefits on one another, the anger system can use threats (and actual implementation) of benefit withdrawal to bargain for a higher WTR. Because of the incredibly cooperative nature of human beings as a species, this is the most common bargaining tactic taken by anger (Averill, 1982). Like aggressive bargaining (see below) cooperative bargaining should begin with low cost demonstrations of one's ability to deny the target benefits, and demonstrations that those benefits are of great value to the target. These demonstrations can include: verbal persuasion in which the angry individual activates memories in the target that contain information about past benefit conferral (e.g. "after I helped you pass the bar you're going to treat me like this!"), conveying hypotheticals about future benefits that can be denied (e.g. "we'll see how you feel the next time you need a ride to work"), deny the potency or frequency of benefits coming to the angry person from the target (e.g. "I don't need you", "fine, I'll just get Bob to take me to work"), and so on.

With those memory evocation and threat games aside, cooperative value is difficult to demonstrate in the short term. This is unlike one's potential to deploy violence; upper body strength, for example, is a component of fighting ability that can be reasonably indicated with a single push. But one's willingness to look after another individual's child is a harder thing to demonstrate. For this reason, demonstrations of cooperative value are typically drawn-out over long periods of time, during which the angry individual adopts a temporarily low WTR toward the target so that the other individual can assess the consequences of cooperative withdrawal. I call these "*contests of withdrawal*." The actual psychological weapons in these contests are demonstrations that the angry person is now valuing the target less (i.e. lowering their own WTR toward the target of anger); such demonstrations can come from any of the usual cues of one's WTR. The signals of lowered WTR, each of which is also, predictably, a trigger of anger, include: restricting information sharing, removing communicative affect from the voice and face (i.e. the silent treatment), ignoring the wants and needs of the target, deriving no apparent pleasure from the target's company, forgetting information about the target, estimating the target's WTR-setting traits as lower than before, feeling less empathically for the target when they experience benefits or costs, and, of course, putting less weight on the target's welfare when making cost-benefit decisions.

Whether this withdrawal of welfare consideration is permanent or merely a bluff appears to be an open variable that is negotiated during the contest of withdrawal. Usually these conflicts can be ended when one individual yields and agrees, implicitly, to treat the individual with more consideration. If that does not happen, the loss of cooperation can be permanent. These contests, like their violent counterparts discussed below, appear to play by certain rules that minimize the long term damage to both contestants and maintain the accuracy of what is being signaled. Just as there is a "fair-fight" for violent conflicts, there is a sense of fairness with conflicts of withdrawal. For example, if two roommates are engaged in an angry conflict and then one discovers that her father was killed in a smelting accident, the other roommate would be expected to immediately resume cooperation. If she were to fail to do this, it would indicate clearly that she was not engaged in a contest of withdrawal but had, in fact, permanently withdrawn from the relationship. This type of situation reveals the temporary bargaining

nature of the contests. There is also reason to believe that conflicts such as this are designed to evoke a "fair" assessment of cooperative value. "Fair," in this context, means an accurate unbiased index of the likely future cooperative value of the two contestants to each other. A conflict would be biased if it occurred when one individual was at a temporary disadvantage that limited their current cooperative value. For example, if a husband waited until his wife were in the hospital before getting angry that she still talks to her ex-husband via text messages, this would be seen by the wife and others as an "unfair" attempt to extract a better WTR than he is otherwise entitled to. It is, in fact, the avoidance of these kinds of temporary fluctuations in bargaining power that characterizes longer term friendships (Tooby & Cosmides, 1996). Such an attempt at bargaining when the target is in a position of temporary weakness indicates that the husband does not value the long term cooperative value of his wife but is only trying to extract as many resources as possible from her in the short term.

2.3.2. Negotiating with cost infliction

Humans, like other animals, respond to relative formidability (i.e. resource holding power, RHP, or the ability to impose costs on others) in ways predictable by basic evolutionary models of conflict. In particular, animals cede more resources to those that are better fighters (Huntingford & Turner, 1987) because the costs of not doing so would be substantial. Viewed in terms of WTRs, when the costs of raising one's WTR toward an angry individual are less than the costs that angry person will impose if you maintain your current WTR, then the rational response is to set a higher weight on the angry individual's welfare. In this way, individuals can use their ability to impose costs on others as a bargaining tactic to raise another's welfare tradeoff ratio. Consequently, the mind should be designed to assess relative formidability from whatever cues were reliably available (Sell et al., 2009; Sell et al., 2010), and respond spontaneously by raising their WTR toward the formidable individual. Furthermore, as with those who can efficiently deliver benefits, individuals who have the potential to inflict great costs on others are predicted, and have been shown, to expect better treatment (i.e. demand a higher WTR) and set a lower threshold for anger (Sell, Tooby & Cosmides, 2009).

As with negotiating with benefit withdrawal, negotiating with formidability ought to proceed in an efficient manner that begins with low-cost displays and arguments and only escalates to the expensive or dangerous demonstrations if those early attempts fail. Formidability, like cooperative value, can come from several channels including personal fighting ability, the size and strength of one's coalitional allies, and social information manipulation. Demonstrating one's formidability through coalitional or social channels will be difficult to do in a short period of time as friends and allies will have to be gathered – allies who may not feel that bargaining for a higher WTR toward the angry person is worth the effort. Still, in the absence of actually imposing costs on the target, verbal tactics are available to threaten or remind the target of personal and coalitional formidability including: declarations of personal fighting ability, e.g. "I can kick your ass", coalitional formidability, e.g. "my dad can beat up your dad", information about reputational damage, e.g. "you better show me some respect unless you want everyone to know what you did in the Hamptons last summer", denigrating the target's formidability, e.g. "you think I'm going to take that from a pussy like you?", and demonstrating resistance to costs from the target, e.g. "there's nothing you can do about it".

These verbal arguments are relatively easy to fake and, for the most part, not informative when negotiating with strangers or individuals with whom one does not have a prior history of negotiating, though they do signal to the target that one is willing to negotiate with cost infliction. If verbal arguments do not recalibrate the target, relative bargaining power may be demonstrated more directly. When it comes to personal fighting ability, selection would design humans like it has other animals, to gather information about

relative fighting ability while minimizing the risks inherent in aggressive combat. Many animals engage in “conflicts of assessment” which have multiple design features that allow for low-cost estimation of fighting ability (Enquist & Leimar, 1983; for review see Huntingford & Turner, 1987).

Firstly, conflicts of assessment typically have a demarcated beginning, such as a stance, vocalization or body posture that exaggerates the aggressor's size and announces the aggressor's intent to bargain with the target. This awards the target a chance to immediately relinquish the contested resource and accept their new negotiated status, or refuse and escalate the conflict.

Secondly, these assessment games are designed to follow a general pattern of escalation in which low cost “ritualistic” demonstrations are exchanged before higher cost attacks or eventual “no holds barred” combat. The logic behind the escalating nature of the conflicts is that large discrepancies in fighting ability should be evident even by comparatively inaccurate demonstrations of fighting ability, and so neither the victor nor the loser will be forced to pay the high costs of violent aggression if a low-cost demonstration reveals the eventual victor. For example, tail beating in the cichlid fish (i.e. shaking one's tail so that it generates waves commensurate with body size and strength) is less predictive than mouth-locked wrestling (during which fish lock mouths and vigorously shake each other), but mouth-locked wrestling is more costly in terms of energy and probability of injury. This pattern of escalation, a model of which is called the sequential assessment game (Enquist & Leimar, 1983), has been observed in numerous and distally-related species including cichlid fish, bowl and doily spider, African buffalo, beetles, common toads, red deer and pigs (see Sell, 2012).

Finally, natural selection appears to have designed specific morphological features or expressions that serve as clear signals of a cessation of aggression on behalf of the organism that admits defeat and relinquishes the contested resource or acknowledges the victor's superior aggressive formidability. Such submission displays are common in a variety of species and are found almost everywhere in the primate order (Smuts et al., 1987).

Human conflicts of assessment look very much like those of other animals. Typical bouts of aggression between adult males fighting over “status” and “face” progress as follows: a verbal or physical signal of challenge is given, the conflict escalates from words to low-cost demonstrations (e.g. staring contests, pushing) and then to more violent aggression (e.g. wrestling, punching, weapon usage), until a signal of surrender ceases the contest or one of the contestants is incapacitated or killed. These same patterns have been found repeatedly in studies of homicides (Berg & Fox, 1947; Daly & Wilson, 1988; Luckenbill, 1977) and assaults (Felson, 1982).

In this view, these violent conflicts are contests to see who can physically best their opponent. There is, up until the point where lethal violence is being deployed, a cooperative aspect to fights that appears to be intuitive even to children, who frequently speak of “fair fights” in which both individuals are in roughly the same tactical position (e.g. neither asleep, with their back turned, distracted). From the standpoint of most theories of aggression, it is not at all clear why some subset of aggressive acts should be labeled “fair” and others “unfair.” If, on the other hand, “fair fights” are evolutionarily derived from the same selection pressures as the sequential assessment games that characterize non-human conspecific aggression then a definition and explanation presents itself: “fair fights” are fights in which human observers accept that accurate information is gleaned regarding individuals' relative fighting ability. This is presumably why the term “unfair” or “cheap” applies to many seemingly different situations such as: using attacks that were implicitly agreed to be off-limits (e.g. hitting in the throat, eye-gouging, strikes to the genitals), fighting someone who is temporarily of lower fighting ability (e.g. hitting someone when their back is turned), and also to fights in which there is a gross difference in fighting ability (e.g. picking on someone half

your size, wrestling a woman ala Andy Kaufman). In the first case, if one individual is using “unfair” attacks while the other had intentionally not deployed those tactics, the result will be biased in favor of the “cheater” for the current contest but not future ones in which both individuals avail themselves of more violent tactics. In the second case, the result does not predict who would win a future fight because the target would be just as likely as the attacker to be able to hit the other from behind – note, however, that if what is being assessed is the ability to move undetected and surprise an opponent then the mind should reclassify a hit in the back as fair. In the third case, using physical aggression against people who have little or no chance of winning the conflict can be seen as attempting to inflate one's reputation inaccurately, e.g. winning ten fights doesn't mean you are a great fighter if you're attacking children.

What is considered fair seems to differ somewhat from culture to culture such that shooting someone in the back was considered cowardly in the “Old West” of America and an individual who did such a thing would not be given the respect earned by a “fair” shooting (Williams, 1980), while a Yanamamo man can sneak into an enemy village and shoot an elderly woman in the back and still be awarded unokai status (Chagnon, 1983). There is reason to believe these differing standards of “fairness” are not solely the result of arbitrary inherited cultural norms, but may be calibrated by the ecological circumstances surrounding combat in that culture. Among the Yanamamo, for example, the ability to traverse the jungle, identify the foe in darkness, fire an accurate shot, escape into the jungle with enemies in pursuit and arrive safely back at camp is an indicator of one's competence at the raiding-type warfare that typifies Yanamamo aggression, even if the target felled by the arrow is not an able-bodied man.

As with assessing cooperative value, a person who engages in “unfair” fights is seen as holding different motives than a person engaging in a conflict of assessment. Conflicts of assessment are used to recalibrate a target's estimation of one's fighting ability and therefore their WTR. This can be distinguished from aggression that is designed to actually kill or damage the opponent. Such aggression is presumably deployed by a different computational mechanism that can be called “hatred.” One theoretical function of hatred is to impose costs on a target in order to reduce their power (either by killing them or weakening them) or incentivize them to avoid interacting with the hateful person. As such, it responds to different input conditions than anger (i.e. you can hate your ex-wife's new husband even if he doesn't know you exist and has never indicated that he has a low WTR toward you). Hatred-based aggression can be distinguished from conflicts of assessment in that conflicts of assessment, i) cease when the individual recalibrates and apologies or concedes defeat, ii) respect the “rules” of fairness, iii) are between individuals of similar fighting ability, iv) are mutual decisions in that neither “backs down”, and v) use an escalating level of violence such that some attacks are implicitly agreed to not be used (e.g. no eye-gouging). Any deviation from those constraints indicates that the attacker is not solely trying to bargain for better treatment but also attempting to impose costs on the target in order to cripple, kill, weaken or incentivize them to leave.

2.3.3. *Choosing among negotiating tactics*

The decision about whether to negotiate by threatening costs or withholding benefits should be a functional one based on the probability of success and efficiency of deployment for each method given the individual one is trying to recalibrate. Thus a number of factors become relevant to this decision, some of which are specific to the angry individuals (e.g. their coalitional and personal formidability) and others specific to the relationship itself (e.g. arguing with a stranger versus arguing with someone who depends on you for their livelihood). In general, it is predicted that an individual will use whichever negotiative tool allows for more bargaining power, though there are complications arising from the fact that formidability-based WTRs

apply to a more narrow range of circumstances than cooperative-based WTRs (Petersen et al., *under review*).³

This simple premise predicts that individuals that are highly skilled at delivering benefits should rely on benefit withdrawal as a negotiative tactic while those who are highly skilled at inflicting costs will be more likely to bargain with aggression and other forms of cost imposition. This prediction fits very nicely with well-established sex and age differences in aggression, in particular, the fact that violent and homicidal aggression are (across all tested cultures and time periods) most common among young men (Daly & Wilson, 1988). Young men have had relatively little time to acquire the complicated skill sets necessary to be the best providers (Gurven, Kaplan, & Gutierrez, 2006) but can rely on their physical strength to bargain for higher WTRs from others.⁴ Even within the class of young men, evidence shows that those who are more efficient fighters – operationalized by upper body strength – expect better treatment from others (demand higher WTRs), set lower thresholds for anger, and are more likely to respond with physical aggression and to believe that physical aggression is an efficacious way to resolve conflicts (Sell, Tooby & Cosmides, 2009). Women will, compared to men, be much less efficient at inflicting physical damage on others and will thus be more likely to rely on indirect aggression or cooperative bargaining (Campbell, 2002).

Individuals who are particularly low in cooperative value should also be more likely to use cost infliction to bargain for better WTRs. This is one explanation for the negative correlations between criminal offending and physical attractiveness (Mocan & Tekin, 2010),⁵ intelligence (Lipsitt, Buka & Lipsitt, 1990) and gainful employment (Williams, Cullen & Wright, 1996). Furthermore, any dramatic decrease in one's cooperative value to another individual should predict a shift toward more cost-infliction type bargaining. The logic is as follows, if an individual no longer values you for your cooperative ability, they will begin to ratchet down their WTR toward you. To prevent this, their WTR will need to be propped up from another source, specifically estimates of your ability to impose costs on them. Indeed, studies on college students and the general public show cooperative rejection leads to aggressive bargaining (Leary, Twenge & Quinlivan, 2006). This tendency to engage in aggression when cooperative routes to status are blocked could explain why mass killers are frequently shown to have been recently fired or rejected by peer groups (Meloy et al., 2004).

Characteristics of the target of anger are also predicted to influence the decision about whether to bargain using cooperative value or formidability. If one is bargaining by using cost-infliction, the angry individual generally forfeits whatever benefits she would have gained by cooperating with the target.⁶ This means that angry individuals

should be averse to using cost-infliction when bargaining with people who are of high potential cooperative value, a prediction borne out by the data (Petersen, Sell, Cosmides & Tooby, *under review*). Additionally, targets who are high in formidability will be difficult to recalibrate using cost infliction, and so angry individuals should find themselves bargaining with them using benefit withdrawal when feasible.

3. A psychologist's guide to eliciting violent anger⁷

As can be seen from the sketch above, violent anger is the result of a long list of computations each of which allows for interventions to short-circuit aggressive action. Starting at the beginning, and illustrated with extreme examples that make each point plainly intuitive, an incident of violent anger must, under normal circumstances and with normal populations, follow this pattern:

1. Initial indication of a low welfare tradeoff ratio toward the target. Anger must be triggered by some window into the mind of the target which allows the angry individual to intuit that the target does not value their interests highly enough. Again, this conclusion can result from many clues.

To efficiently trigger anger from cost-inflictions one needs to demonstrate the lowest possible WTR, in other words, impose large costs on the target for trivial benefits with full knowledge of the identity of the person on whom that cost was imposed. In fact, shared attention with the target will also make it clear that you are processing the identity of the person, recognize that you are imposing a cost on them, and continuing to do it anyway. As long as the ratio of cost to benefit is large, the absolute magnitude of the cost imposed need not be high. For example, if a colleague takes off his sweater at lunch you can simply take a sleeve – look the owner in the eye – and wipe some ketchup off of your face with it.

To trigger violent anger with insults, one should perform a clear, unambiguous act that demonstrates both that an individual is deficient on a WTR-relevant trait and that because of that deficiency, you value their welfare less. This can be done very directly, e.g. “I think your theory is worthless and I question your suitability for academic employment. Having read the paper, I believe I'll be spending less time speaking with you.” To maximize the probability of a violent response however, the insult should be on a trait relating to their formidability, e.g. “What are you going to do about it, wimp?” or “You're a pushover, and if you did absolutely everything in your power to harm me, I wonder if I'd even notice?”

Importantly, the indication of a low WTR must be credible to the angry individual given the nature of their relationship and previous evidence. For example, if a very gentle and kind friend wiped their mouth with your sweater, you might be more likely to conclude that they were having a bad day or suffering from a mental disease than that they suddenly valued you less.

2. Stage 1 anger must escalate. Once anger is triggered it will usually deploy search algorithms in the angry person – indeed when angry it is difficult to think or talk about anything else. These are designed to test whether the triggering event is an accurate representation of the target's WTR. Additional cues that cement the angry individual's estimate of one's low WTR will rapidly increase the chance of provoking violent anger.

³ It is likely that individuals have at least two distinct WTRs toward others, an intrinsic WTR set largely by cooperative value, kinship and mutual dependence, which is deployed even in the absence of the individual, and a monitored WTR set largely by direct reciprocity and formidability that is used only when the individual can defend his or her interests. One benefit of a cooperative partner is that they have an incentive to consider your interests even when you are not around to monitor them, i.e. their WTR remains high in your absence because they have an incentive to keep you alive, healthy and functioning in your cooperative capacity. This point is illustrated by the fact that it was not at all uncommon to find slaves, from various cultures around the world, paid for their services because masters discovered that they could extract more economic effort by cultivating cooperative interactions rather than relationships based on cost-infliction.

⁴ This should not be taken to suggest that the only reason young men are particularly violent is because they have little in the way of cultivated skills and more in the way of the ability to inflict costs. There are many other reasons why young males may be more violent (Daly & Wilson, 1988).

⁵ This relationship is complicated, in men, by the fact that attractiveness and physical strength are highly correlated (Sell, Tooby & Cosmides, 2009, SI).

⁶ This is because WTRs that are maintained by the ability to impose costs do not incentivize the individual to benefit the other in their absence. For example, when the bully is in the room a child will treat him with deference, but they are not likely to invite them to parties or otherwise go out of their way to engage in cooperative action. Indeed, high WTRs set by cost-infliction turn the other individual into a noxious stimulus whose presence predicts costs both from bargaining and from having to consider their welfare when making decisions. See footnote 1 as well.

⁷ When explaining this theory to others, I've found more skepticism when listing variables that can assuage violent anger than when listing those same variables (with reverse polarity) as exacerbating violent anger. For example, people are skeptical that violence is less likely when a person imposes a cost on you at random (e.g. “I didn't know it was your car I was keying!”). The typical response is, “that would still make me angry.” But that same individual is not skeptical that a violent response would be more likely if the person imposed the cost on him in particular (e.g. “I knew it was your car, that's why I keyed it!”). For that reason, I find it easier to show the exacerbating effect of each variable on violent anger.

This is particularly true of cost inflictions which make up the bulk of human anger-based arguments. In order to have an accurate estimate of a target's WTR from their imposition of a cost, one requires information about how large of a cost they anticipated it would be, how large of a benefit they anticipated receiving, that they knew the identity of the person on whom the cost would be imposed, and that they weighed the cost/benefit ratio prior to the act. To the extent that the act or subsequent statements about the act indicate knowledge of these variables, one can hasten the conclusion in the angry individual's mind that one does not value their welfare highly, and escalate to stage 2 anger. For example, if your colleague – in response to you wiping ketchup on his sweater – confronted you with what you've done and said, "You know you just wiped ketchup on my sweater, right?!" you could respond with a simple, "Yes."

Furthermore, because these search routines in the angry person are generally meant to catch false positives (i.e. inappropriate anger at a small offense) the escalation to stage 2 anger can be hastened by indicating that the angry person did misread your estimate of the costs and benefits, but in the *other* direction, i.e. the target actually thought the cost was *higher* and the benefit *lower*. For example, in the case of the ketchup and the sweater, after wiping the ketchup from your face you can point out that you thought you had some mustard on your mouth as well, and at first you thought it was that sweater his son bought him for Christmas before he was killed in that car accident. Furthermore, to confirm that you know the identity of the individual on whom you are imposing the cost you can ask the owner of the sweater beforehand if that is indeed their sweater, and when they confirm that it is, proceed to wipe the ketchup off your mouth. This last offense is doubly anger provoking because it confirms you are willing to impose the cost *on him* but also because it implies that you would not have done it to others. Comparisons of this kind are very common in anger arguments as they hold the cost/benefit ratio constant and vary only the WTR, allowing for accurate information about how much someone values you relative to another, e.g. "I wouldn't do that to you", "You let Bob borrow your bike but not me?"

In the case of anger triggered by insults or insulting beliefs, one can lessen anger by pointing out compensatory factors that override the disparaged factor and lead to a high WTR, e.g. "John isn't very smart, but he's the most loyal friend I've ever known and I'd die for him." Likewise, to exacerbate anger one can indicate that the target is deficient on multiple factors, e.g. "Not only are you a bad science writer, but I find your personality repulsive, your courage suspect, and you are known to emit a foul and unpleasant odor that is distracting to your colleagues." Insults regarding traits that are relevant to cost-infliction should be particularly likely to trigger violence, e.g. personal fighting ability, courage.

Regardless of the trigger, convergent evidence of a low WTR becomes more convincing. Anyone can mistakenly give the impression of not valuing someone once in a while, but over the long term a person who genuinely devalues your interests will radiate cues repeatedly. For this reason, repeated indicators of a low WTR will be much more likely to trigger and escalate anger to stage 2.

Repeated impositions of costs are also likely to escalate anger because they bear on the cost of ignoring the low WTR. A low WTR can be cost-effectively ignored if the individual who holds it does not interact frequently with the person held in low esteem. To weight the decision in favor of escalation, you should convince the angry individual that you fully intend to spend a great deal of time with them, judging them, imposing costs on them, and sharing your low opinions of their WTR-relevant traits with those in their vicinity.

3. The target resists low cost recalibrations. With few exceptions, stage 2 anger starts with low cost recalibrations when possible and only escalates to violence after those have failed. Indications of this "failure" can be generated in a number of ways. For example, when an individual is angry with you, you can completely ignore their anger. This isn't to say one should be quiet or avoid them, those are common

responses consistent with reevaluating what you have done; but rather be uninterested in their anger and go on about what you were doing before, e.g. "I can tell that you're angry; I don't care, please pass the salt."

If the angry person is using threats of physical aggression in their early negotiations (i.e. posturing their body, making a closed-mouth anger expression, yelling), then escalation can be achieved by showing no fear and seeming relaxed and unimpressed with their physical strength. Laughing and pointing at the angry individual will often be sufficient to indicate that you are not prepared to elevate your WTR in response to their threats of physical harm, indeed you believe their ability to harm you at all is suspect, pulling an additional trigger of anger. If the angry person is negotiating by withdrawing benefits, then similarly indicating that one is not concerned about any decrease in benefits or that you are positively relieved at the prospect of losing the cooperative relationship, e.g. "I can't hang out with you anymore? I guess I'll just have to relax and have fun instead."

Ordinarily, anger at the early bargaining stage can be immediately diffused with a genuine apology, i.e. a declaration that the target has recalibrated and will, in the future or currently, weight the angry individual's welfare more highly. In fact, this is probably the most effective way for a target to diffuse anger (Frantz & Bennigson, 2005). It stands to reason then, that "anti-apologies" should increase the prospect of violent anger. The English language lacks a word for an anti-apology, though a consultation of the etymology of the word would suggest as its opposite something called an "emlogy," the issuing of which can be called "emlogizing." This is the rejection of anger by the target and the claim that the target not only acted appropriately but that had the affront indicated an even lower WTR that would also be acceptable. It is a kind of reflexive anger that says, "how dare you get angry with me?!" Claims of this sort indicate rather clearly that there is a gulf between the WTR the angry person feels entitled to and the WTR the target holds for the angry person. Emlogizing, in practice, simply means deploying anger against the individual who is angry with you. To continue the ketchup/sweater example, if your colleague demands an apology for ruining his sweater, an emlogy would be something like, "No chance in hell! You owe me an apology for being such an ass about this."

4. Angry individual uses cost-infliction rather than benefit-withdrawal to bargain. Once anger has been activated and opportunities for early recalibration have been unsuccessful, the system has two primary tools with which to recalibrate the target: the imposition of costs and the withdrawal of benefits. The anger system is predicted to deploy whichever tactic would most efficiently recalibrate the target. This decision is influenced by many factors that cannot easily be modified in the short term; for example, when dealing with strangers with whom one has no previous cooperative relationship, negotiating with cooperative withdrawal is a less attractive option.

To maximize the chances that the angry person will resort to violence, one has to indicate that their efforts at negotiating with benefit withdrawal are futile, e.g. "I don't care if you're mad and not talking to me." One can also begin to bargain with cost imposition yourself. If someone says they will not be your friend, a strong push or shove may convince them that their attempts at withdrawal-based bargaining didn't work. Of course, if negotiating with benefit-withdrawal is ineffective, the angry person may simply withdraw from the relationship and cease bargaining. To evoke violent anger in them, one has to revisit step two and reconfirm that the low WTR toward them cannot be efficiently ignored.

4. Conclusions

The disordering of other organisms – aggression – is a tool that animals have been designed to avail themselves of when, over evolutionary history, doing so led to more reproductive success. This tool is available to multiple mental systems such as hatred, play,

predation, predator defense, warfare, fear and anger. To understand aggression deployed by any of these systems we require a computational account of the system itself, what triggers it, what variables it responds to, and what behavioral tools it deploys before aggression. The recalibrational theory is a cognitive-evolutionary model of how human anger works based on its putative function: the negotiation of better welfare tradeoff ratios.

According to this model, anger is a computational system that comes online when it detects cues of a low WTR in others, i.e. cues that the other individual does not weight the angry individual's welfare highly when making decisions that impact them both. These cues can come from a variety of mental programs; the most theoretically clear is a cost imposition in which a large cost is imposed on the angry person by someone who gains little from it. In response to these cues the anger system will gather information about the infraction to determine the actor's actual WTR and then, if it is sufficiently low, deploy tactics to recalibrate the target's WTR upward. These tactics include benefit withdrawal and cost infliction. According to the recalibration theory, aggression is normally deployed by the anger system only after, i) the target of anger gives an indication that they hold a low WTR toward the angry individual, ii) early information searches provide convergent evidence that the WTR is too low, iii) the target resists early, low cost attempts at recalibration and finally iv) negotiating with cost infliction is deemed, by the anger system, as more efficient than negotiating with benefit withdrawal.

References

- Archer, J. (1988). *The behavioural biology of aggression*. Cambridge: Cambridge University Press.
- Averill, J. (1982). *Anger and aggression: An essay on emotion*. New York: Springer-Verlag.
- Berg, I., & Fox, V. (1947). Factors in homicides committed by 200 males. *Journal of Social Psychology*, 26, 109–119.
- Bugental, D., Blue, J., & Cruzcosa, M. (1989). Perceived control over caregiving outcomes: Implications for child abuse. *Developmental psychology*, 25(4), 532–539.
- Campbell, A. (2002). *A mind of her own: The evolutionary psychology of women*. London: Oxford University Press.
- Chagnon, N. (1983). *Yanomamo: The fierce people* (3rd ed.). NY: Holt, Rinehart and Winston.
- Daly, M., & Wilson, M. (1988). *Homicide*. New York: de Gruyter.
- Duntley, J., & Buss, D. (2004). The plausibility of adaptations for homicide. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The structure of the innate mind*. New York: Oxford University Press.
- Ekman, P., Levenson, R., & Friesen, W. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, 221(4616), 1208–1210.
- Enquist, M., & Leimar, O. (1983). Evolution of fighting behaviour; decision rules and assessment of relative strength. *Journal of Theoretical Biology*, 102, 387–410.
- Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *The American Economic Review*, 90(4), 980–994.
- Felson, R. (1982). Impression management and the escalation of aggression and violence. *Social Psychology Quarterly*, 45(4), 245–254.
- Frank, R. H. (1988). *Passions within reason: The strategic role of the emotions*. New York: Norton.
- Frantz, C., & Bennis, C. (2005). Better late than early: The influence of timing on apology effectiveness. *Journal of Experimental Social Psychology*, 41(2), 201–207.
- Geen, R. (1998). Aggression and antisocial behavior. In D. Gilbert, & S. Fiske (Eds.), (4th ed). *The handbook of social psychology*, Vol. 2. New York, NY, US: McGraw-Hill.
- Curven, M., Kaplan, H., & Gutierrez, M. (2006). How long does it take to become a proficient hunter? Implications for the evolution of extended development and long life span. *Journal of human evolution*, 51(5), 454–470.
- Hansen (1986). Fighting behavior in bald eagles: A test of game theory. *Ecology*, 67(3), 787–797.
- Harmon-Jones, E., & Sigelman, J. (2001). State anger and prefrontal brain activity: Evidence that insult-related relative left-prefrontal activation is associated with experienced anger and aggression. *Journal of Personality and Social Psychology*, 80(5), 797–803.
- Harris, M. B. (1993). How provoking: What makes men and women angry? *Aggressive behavior*, 19(3), 199–211.
- Hirshleifer, J. (1987). On the emotions as guarantors of threats and promises. In J. Dupré (Ed.), *The latest on the best: Essays on evolution and optimality* (pp. 307–326). Boston: MIT Press.
- Huntingford, F. A., & Turner, A. K. (1987). *Animal conflict*. New York: Chapman & Hall.
- Lazarus, R. (1991). *Emotion and adaptation*. Oxford: Oxford University Press.
- Leary, M., Twenge, J., & Quinlivan, E. (2006). Interpersonal rejection as a determinant of anger and aggression. *Pers soc psychology rev*, 10(2), 111–132.
- Lipsitt, P., Buka, S., & Lipsitt, L. (1990). Early intelligence scores and subsequent delinquency. *American journal of family therapy*, 18, 197–208.
- Luckenbill, D. (1977). Criminal homicide as a situated transaction. *Social Problems*, 25, 176–186.
- Meloy, J. R., Hempel, A. G., Gray, B. T., Mohandie, K., Shiva, A. and Richards, T. C. (2004). A comparative analysis of North American adolescent and adult mass murderers. *Behav. Sci. Law*, 22, 291–309.
- Mocan, N., & Tekin, E. (2010). Ugly criminals. *The review of economics and statistics*, 92(1), 15–30.
- Moyer, K. E. (1968). Kinds of aggression and their physiological basis. *Communication in behavioral biology*, 2(2), 65–87.
- Ohbuchi, K., & Kambara, T. (1985). Attacker's intent and awareness of outcome, impression management, and retaliation. *Journal of Experimental Social Psychology*, 21, 321–330.
- Petersen, M., Sell, A., Cosmides, L. & J. Tooby (under review). To punish or repair? Evolutionary psychology and public opinion toward criminal justice.
- Preston, K., & Kimberley, S. (1987). "What's the worst thing...?" Gender-directed insults. *Sex Roles*, 17(3), 209–219.
- Sell, A. (2012). Evolved cognitive specializations. In J. Vonk, & T. K. Shackelford (Eds.), *The Oxford handbook of comparative evolutionary psychology*. New York: Oxford University Press.
- Sell, A., Cosmides, L., Tooby, J., Lim, J., Sugiyama, L. & Krauss, A. Forthcoming. Anger and welfare trade-off ratios: Mapping the computational architecture of a recalibrational emotion system.
- Sell, A., Bryant, G., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., et al. (2010). Adaptations in humans for assessing physical strength and fighting ability from the voice. *Proceedings of the Royal Society B*, 277, 3509–3518.
- Sell, A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., & Gurven, M. (2009). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings of the Royal Society B*, 276, 575–584.
- Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger. *Proceedings of the National Academy of Science*, 106(35), 15073–15078.
- Sherman, P. (1981). Reproductive competition and infanticide in Belding's ground squirrels and other organisms. In R. D. Alexander, & D. W. Tinkle (Eds.), *Natural selection and social behavior: Recent research and new theory* (pp. 311–331). New York: Chiron.
- Smuts, B., Cheney, D., Seyfarth, R., Wrangham, R., & Struhsaker, T. (1987). *Primate societies*. Chicago: The University of Chicago Press.
- Swain, J. (1998). What does bullying really mean? *Educational Research*, 40(3), 358–364.
- Taylor, S. (1967). Aggressive behavior and physiological arousal as a function of provocation and the tendency to inhibit aggression. *Journal of Personality*, 35(2), 297–310.
- Tooby, J., & Cosmides, L. (1996). Friendship and the Banker's paradox: Other pathways to the evolution of adaptations for altruism. *Proceedings of the British Academy*, 88, 119–143.
- Tooby, J., Cosmides, L., Sell, A., Liberman, D., & Sznycer, D. (2008). Internal regulatory variables and the design of human motivation: A computational and evolutionary approach. In A. Elliot (Ed.), *Handbook of approach and avoidance motivation*. Mahwah, NJ: Lawrence Erlbaum Associates.
- von Rueden, C., Curven, M., & Kaplan, H. (2008). The multiple dimensions of male social status in an Amazonian society. *Evolution and Human Behavior*, 29, 402–415.
- White, G. F. (1975). Public responses to hypothetical crimes: Effect of offender status and seriousness of the offense on punitive reactions. *Social Forces*, 53, 411–419.
- Williams, J. (1980). *Duelling in the old south. Vignettes of social history*. College Station, TX: Texas A & M University Press.
- Williams, N., Cullen, F. T., & Wright, J. P. (1996). Labor market participation and youth crime. *Social pathology*, 2, 195–217.