

UNIVERSITY OF CALIFORNIA

Santa Barbara

Regulating Welfare Tradeoff Ratios: Three Tests of an Evolutionary-Computational  
Model of Human Anger

A Dissertation submitted in partial satisfaction of the  
requirements for the degree Doctor of Philosophy  
in Psychology

by

Aaron Sell

Committee in charge:

Professor Leda Cosmides, Chair

Professor John Tooby

Professor Daphne Bugental

Professor James Roney

September 2005

The dissertation of Aaron Sell is approved.

---

John Tooby

---

Daphne Bugental

---

James Roney

---

Leda Cosmides, Committee Chair

August 2005

Regulating Welfare Tradeoff Ratios: three tests of an evolutionary-computational  
model of human anger

Copyright © 2005  
by  
Aaron Sell

Dedicated to those who stay on the bus

## ACKNOWLEDGEMENTS

I'd like to thank my advisor Leda Cosmides. Leda, I admire you for your passion and your bravery in the pursuit of truth. I want to thank you for always being on my side, for your patience, for your encouragement and the excited look you always get on your face whenever I report a new data point. You are a brilliant and shining example of all that is righteous and good in academia. I'd also like to thank John Tooby. John, the reason you can't convince anyone that the world is as depressing and corrupt as you see it is because we all live in a world with you in it. You're the wisest man I've ever met, but you've never made me feel stupid.

Thank you both. You've made me as smart as I always thought I already was.

Finally, I'd like to gratefully acknowledge financial support for this research from the UCSB Academic Senate and a research gift from Mr. Howard Waldow.

## VITA OF AARON SELL

August 2005

### EDUCATION

Bachelor of Arts in Psychology, Ohio University, Athens, May 1995 (summa cum laude)

Doctor of Philosophy in Psychology, University of California, Santa Barbara, August 2005

### PROFESSIONAL EMPLOYMENT

1996-1997: Teaching Assistant, Department of Psychology, Ohio University

1998-2002: Teaching Assistant, Department of Psychology, University of California, Santa Barbara

2003-2004: Teaching Assistant, Department of Communication, University of California, Santa Barbara

Summer 2001, 2002, 2004: Teaching Associate, Department of Psychology, University of California, Santa Barbara

### PUBLICATIONS

Sell, A., Cosmides, L., & Tooby, J. (2002). Evolutionary Psychology: Applications and Criticisms. *Encyclopedia of Cognitive Science*. Nature Publishing Group..

### AWARDS

Dissertation Research Fellowship, UCSB Graduate Division, Summer 2005

Humanities and Social Science Research Grant, UCSB, Fall 2004

GSA Excellence in Teaching Award – Outstanding Teaching Assistant in Social Sciences, UCSB, 2000

UCSB Special Regents Fellowship, UCSB, 1998-2002

### FIELDS OF STUDY

Major Field: Evolutionary Psychology

## ABSTRACT

### Regulating Welfare Tradeoff Ratios: Three Tests of an Evolutionary-Computational Model of Human Anger

by

Aaron Sell

Competition over scarce resources has always been an adaptive problem for organisms, and natural selection has designed efficient, effective, reliably-developing mechanisms for adaptively regulating resource competition in a host of non-human animals. These mechanisms have been shown to compute the relative Resource Holding Power (i.e. fighting ability) of the contestants, the relative value of the resource to the two contestants, and regulate demonstrations of those valuations often in the form of violent contests. These mechanisms store an internal regulatory variable that determines the extent to which an organism will impose costs on another organism for its own gain (i.e. a marker of status that determines dominance hierarchies). Herein it is assumed that human beings also have internal regulatory variables that store the extent to which they are willing to gain benefits by imposing costs on another and the extent to which they are willing to tolerate costs from another. These variables are called Welfare Tradeoff Ratios. It is theorized that one of the primary functions of human anger is to regulate resource conflict by increasing the target's Welfare Tradeoff Ratio toward the angry individual such that the target becomes less willing to impose costs on that individual and more willing to tolerate

costs. Three primary predictions are derived from this model and empirically tested using a range of different methodologies. The predictions are: 1) anger should be triggered by cost-benefit transactions that impose large costs on the angry individual for relatively small benefits; 2) in males, assessments of physical strength should be among the variables that regulate the magnitude of Welfare Tradeoff Ratios; 3) human arguments should be interpretable as attempts to modify perceptions of a cost-benefit transaction in ways that fit the logic of the selection pressures of resource conflict. In all three study sets, predictions derived from the Regulator Theory of Human Anger were experimentally confirmed. Alternative models of human anger are discussed and reviewed with respect to their ability to account for these findings as well as those reported in the prior literature on anger and aggression.

## TABLE OF CONTENTS

Chapter 1. Selection pressures that lead to a cognitive model of human anger.....	1
A. Selection pressures inherent in resource conflict.....	2
B. The design of animal conflict reflects the selection pressures in the Asymmetric war of Attrition.....	6
1. Relative resource holding power.....	6
2. Relative valuation of the resource.....	7
3. Assessment of relative RHP and relative resource value.....	8
4. Relative valuation of the resource.....	13
5. Animals are designed to demonstrate their valuation of the resource.....	15
6. Species will maintain internal representations of relative RHP.....	16
C. Conclusion about animal aggression.....	18
D. A cognitive model of animal aggression.....	18
E. Humans: constructing a cognitive model of human anger.....	20
1. Welfare Tradeoff Ratios (WTRs).....	23
2. Regulator theory of human anger.....	27
3. Summary of regulator theory of human anger.....	33
Chapter 2. Alternative theories and how they explain known features of anger.....	35
A. Known features of anger.....	35
1. Feature #1: anger frequently results from the imposition of costs.....	36
2. Feature #2: intentionality increases the likelihood of anger.....	36

3. Feature #3: apologies mitigate anger.....	37
4. Feature #4: anger and aggression are often used by males to restore “face”.....	38
5. Feature #5: personal insults are one of the most reliable causes of anger...	39
6. Feature #6: anger has cross-culturally universal features and neuro- physiological locality.....	39
B. Alternative theories of anger.....	41
1. Cognitive-Motivational-Relational theory (Richard Lazarus).....	42
2. Cognitive Neoassociationism (Leonard Berkowitz).....	47
3. Equity Theory.....	51
4. The Frustration-Aggression Model.....	54
5. A behaviorist account of anger (Arnold Buss).....	56
6. Social Learning Theory (Bandura).....	58
7. Social Constructivism (James Averill).....	61
8. Social Interactionist Perspective (Tedeschi & Felson).....	64
9. The Regulator Theory of Anger.....	66
C. Summary of theories and their ability to account for features of anger.....	70
Chapter 3. Study Set A – the causes of anger elaborated.....	72
A. Experiment A1: Is less anger elicited when a cost is imposed to obtain a high benefit than when the same cost is imposed to obtain a low benefit?.....	72
B. Experiment A2: Is anger mitigated when the offender did not know the angry individual’s identity?.....	89

C. Experiment A3: Can human responses to anger be understood as attempting to indicate a high Welfare Tradeoff Ratio and thus assuage anger?.....	100
D. Experiment A4: Converging evidence: Which arguments in fact assuage anger?.....	108
E. Discussion of Experiment Set A.....	120
Chapter 4. Individual differences in male anger – Study Set B.....	121
A. Formidability and human anger.....	121
1. Formidability predicts aggression in non-humans.....	121
2. Formidability should predict aggression in human males.....	124
3. More formidable males should view aggression as a more effective way to resolve interpersonal disputes and be less afraid of retaliation.....	125
4. More formidable males should view aggression as a more effective way to resolve international disputes and be less afraid of retaliation in such disputes.....	125
B. Predictions regarding strength and anger from other theories.....	126
1. Cognitive-Motivational-Relational Theory.....	127
2. Cognitive Neoassociationism and Frustration-Aggression Model.....	127
3. Equity Theory.....	128
4. Behaviorism.....	129
5. Social Learning Theory.....	130
6. Social Constructivism.....	130
7. Social Interactionism.....	131

8. “Short Man Syndrome” or “Napoleon Complex” .....	131
C. Study B1: Strength predicts anger in college weight-lifters.....	131
1. Can alternative theories of anger account for these effects?.....	144
2. Pragmatic obstacles for Study B1 .....	150
D. Study B2: Strength predicts anger outside the gym.....	152
E. Study B3: Replication of strength and anger with a larger sample.....	165
F. General Discussion for Study Set B.....	178
Chapter 5. A New direction in the behavioral output of anger – Study C.....	182
A. The function of human arguments according to the Regulator Theory.....	182
1. Recalibrate the target’s Welfare Tradeoff Ratio.....	184
2. Change the target’s estimates of the consequences of their actions.....	186
B. How do other theories account for argumentation?.....	188
1. Cognitive-Motivational-Relational.....	188
2. Cognitive Neoassociationism.....	188
3. Equity Theory.....	189
4. Frustration-Aggression Model.....	189
5. Behaviorism.....	190
6. Social Learning Theory.....	191
7. Social Constructivism.....	192
8. Social Interactionism.....	192
C. Study C – A preliminary analysis of naturalistic human arguments.....	193
Chapter 6. Conclusions and unanswered questions.....	201

A. New features of anger.....	201
1.Feature #7: Anger becomes increasingly more likely to result from a cost-benefit transaction as the benefit to the target of anger decreases....	201
2. Feature #8: Individual differences in the proneness to anger, in males but not females, is correlated with physical strength.....	202
3. Feature #9: A sizable proportion of human arguments are designed to manipulate perceptions of Welfare Tradeoff Ratios.....	202
B. How do other theories account for the new features of anger.....	202
1. Cognitive-Motivational-Relational Theory.....	203
2. Cognitive Neoassociationism.....	205
3. Equity Theory.....	205
4. Frustration-Aggression Model.....	206
5. Behaviorism and Social Learning Theory.....	207
6. Social Constructivism.....	208
7. Social Interactionism.....	208
C. The Regulator Theory accounts for the known features of anger.....	209
1. Feature #1: Larger costs cause more anger than smaller costs.....	209
2. Feature #2: Intentionally imposed costs are more anger provoking.....	210
3. Feature #3: Apologies reduce anger.....	210
4. Feature #4: Anger and aggression are often used by males to restore “face”.....	211
5. Feature #5: Personal insults are one of the most reliable causes of anger.	212

6. Feature #6: Anger has cross-culturally universal features and neuro- physiological locality.....	213
D. Conclusions – how to account for complex order .....	213
References.....	219

## LIST OF FIGURES

Figure 1-1. Examples of cost-benefit transactions as ratios of value.....	23
Figure 1-2. Welfare Tradeoff Ratio as a threshold of cost-benefit transactions.....	25
Figure 1-3. The Basic Computational Process of Anger.....	29
Figure 3-1. Change in anger in high and low benefit conditions.....	81
Figure 3-2. Self-report of emotions when offender received high benefit (HB) or low benefit (LB) as a result of imposing a cost.....	84
Figure 3-3. Change in anger in known and anonymous victim conditions.....	93
Figure 3-4. Self-report of Emotions when offender knew and didn't know the victim's identity.....	95
Figure 3-5. Which arguments to subjects believe will assuage or exacerbate another's anger?.....	105
Figure 3-6. Which arguments do assuage or exacerbate another's anger?.....	111
Figure 3-7. Preference for argument in Experiment A3 (reverse coded) with the argument's impact on anger in Experiment A4.....	116
Figure 4-1. Zone of Contestment for organisms that are higher, equal or lower in formidability than the other contestant.....	123

## **Chapter 1.**

### **Selection Pressures that lead to a Cognitive Model of Human Anger**

Natural selection is the only process shown capable of creating complex functional design in organisms (Williams, 1966). As such, the existence of specialized machinery for solving adaptive problems is strong evidence that natural selection has designed the machinery for that adaptive problem, or for a broader class of adaptive problems. This causal relationship enables two conclusions about the nature of living organisms. One, evidence of strong selection pressures over a significant portion of a species' evolutionary history, coupled with the existence of physiologically possible solutions to said selection pressures,<sup>1</sup> give reason to believe that adaptations for solving that selection pressure have been designed into the species upon which the selection pressures acted. And two, the existence of functionally-designed machinery in an organism is evidence for the ancestral presence of selection pressures for that function. The adaptationist program is an approach to life sciences that uses these tenets to match selection pressures to the mechanisms they have designed (Williams, 1966). This approach allows one to better understand selection pressures by studying design (e.g. discovery of the selection pressures responsible for sexual reproduction were found after Williams closely examined the machinery of sexual recombination [1975]; Tooby 1982), and to

---

<sup>1</sup> Selection pressures can only act on existing genetic variance, and as such are incapable of designing features that require great leaps of complexity before any increase in fitness. For example, there is a selection pressure on human males to instantly impregnate females from miles away, but given that random mutations in the current population do not lead to any incremental chance of solving the problem it is currently impossible for natural selection to design such mechanisms.

discover and understand organic design by studying selection pressures (e.g. adaptive specializations for reasoning about social exchange were not discovered until the selection pressures behind it were mapped [Williams, 1966; Trivers, 1971; Axelrod & Hamilton, 1981; Cosmides, 1989; Cosmides & Tooby, 1992]).

Evolutionary psychology is an approach to psychology that uses analyses of selection pressures to discover and understand the cognitive machinery designed by those selection pressures. The closer the fit between the logic of the selection pressures and the known features of the mechanism under study, the more confident one can be in the proposed function of the mechanism, the accuracy of the posited selection pressures, and any future predictions derived from the model of the machinery under study. Therefore, a thorough examination of both the selection pressures and the proposed cognitive mechanism are necessary for a theoretically-sound, computational model of any organic mechanism designed by natural selection.

#### ***A. Selection pressures inherent in resource conflict***

By virtue of their common design, organisms of the same species will tend to exploit the same resources in their environment; e.g., mate with the same species, eat the same foods, seek the same territories. Selection pressures to gain disproportionate access to finite resources will then design organisms to compete for those resources. Excluding kin, and assuming there are no reciprocal consequences, natural selection will design one organism to attempt to gain resources with no regard for the impact of their activities on the reproduction and survival of competitors. Considering this, the animal kingdom presents a surprisingly high number of species that demonstrate

apparent restraint in response to conflicts of interests. Rattlesnakes, for example, use wrestling to resolve conflicts rather than their deadly fangs (Maier, 1998). Red deer (elk) will parallel walk, roar, and finally antler wrestle but do not stab each other from behind or the side with antlers (Clutton-Brock & Albon, 1979). Cichlid fish engage in parallel swimming, mouth wrestling and tail beating before engaging in highly-damaging combat (Enquist & Leimar, 1983; Enquist, Leimar, Ljungberg, Mallner & Segerdahl, 1990). Evidence of phenotypic design causing restraint was counter to the posited selection pressures and led to a reexamination of the selection pressures responsible for animal conflict.

These examples were so contrary to the perceived logic of natural selection, that group selection was deemed the only plausible explanation (Lorenz, 1966; Wynne-Edwards, 1962), suggesting that these contests were not selected for at the individual level, but rather that genes for engaging in sublethal contests were selected because groups that contained such individuals would survive longer than groups that contained lethal fighters.

In the same year of Lorenz's publication, George Williams published *Adaptation and Natural Selection* (1966), the most comprehensive refutation of then extant versions of group selection theory (see also, Dawkins 1976; Lack, 1966). Lorenz's explanation of restrained conflict (i.e. ritualistic conflict, in his terms) suffered from a number of William's criticisms. Specifically, animals that used lethal strategies would, according to Lorenz, increase in number within all groups so, for there to be a between group advantage to restraint, groups with disproportionately high numbers of

restrained fighters would need to fractionate and reform continuously in order to avoid being composed entirely of lethal fighters. There is no explanation as to how or why such fractionations would occur, let alone why the positive assortment, which would be required to avoid diffusion of ritualistic fighters and the eventual selection of lethal fighters, would take place.

A number of theories of animal conflict arose to fill the vacuum after the dissolution of group selection accounts; most of these new theories stemmed from the introduction of game theory to evolutionary biology by Maynard Smith (1982), which allowed evolutionary biologists to posit mathematically precise models of selection pressures.

The Asymmetric War of Attrition (AWA; also known as the generalized war of attrition) was one of those explanations. It is an economic model for analyzing the selection pressures surrounding situations of conflict between two animals of different strength competing over an adaptive resource that holds differential value to both organisms (Hammerstein & Parker, 1982; Maynard Smith & Parker, 1976). The model expresses the conditions under which X will fight if Y does not relinquish the resource. For any two organisms, X and Y, X should fight for the resource when the following inequality is satisfied:

$$v(X) > v(Y) * k(X)/k(Y)$$

In this equation,  $v(X)$  is the adaptive value of the contested resource to organism X (i.e. the fitness difference, for organism X, between two outcomes: X receiving the resource and X not receiving the resource),  $v(Y)$  is the adaptive value of the resource for organism Y, and  $k(X) / k(Y)$  is relative RHP (i.e. Resource Holding Power or Resource Holding Potential) which is the rate at which organism X will incur injuries if both organisms attempt to gain the resource divided by the rate at which Y will incur injuries. The higher the relative RHP the more damage Y can do to X if both attempt to get the resource. When the equation is satisfied, organism Y will be selected to relinquish the contested resource without a fight and organism X will be selected to engage in the fight if Y does not relinquish the resource.<sup>2</sup>

Considering the effect of each variable in the AWA will give an intuitive feel for the formula. Imagine you (organism X) and organism Y happen to find a morsel of food on the ground. The model shows that you will be selected to attempt to take the resource if:  $v(X)$  is high (you will benefit a lot from the food),  $v(Y)$  is low (Y will benefit little from the food), and relative RHP is low (you are “tougher” than Y).

The logic behind the model is as follows. Organisms will be selected to attempt to gain a resource if the fitness benefit of obtaining the resource is greater than the fitness cost of acquisition. The benefit of obtaining the resource will be determined by the value of the resource itself (i.e. the left side of the equation). The cost of acquisition will be determined by the relative fighting ability between the two

---

<sup>2</sup> More complicated versions of the model make more complicated predictions, but the role of the variables stays the same: the more an organism values a resource, the more likely it is to receive it, and the more formidable the organism is the more likely it is to get the resource.

organisms and value to the competitor (i.e. the right side of the equation). The value of the resource to the organisms becomes relevant to the resolution of resource conflict because it corresponds to the fitness budget that an organism will be selected to spend to get the resource – for example, an organism will be selected to expend 5 reproductive units to get a resource worth 6 reproductive units, but not 7 to get that same resource. It is because of the differential fitness budget (i.e.  $v(X) / v(Y)$ ) that less formidable animals are capable, at times, of getting resources valued by more formidable opponents.

***B. The design of animal conflict reflects the selection pressures in the Asymmetric War of Attrition***

Six converging lines of evidence suggest that animals have been designed to adaptively respond to the selection pressures modeled in the Asymmetric War of Attrition.

1. Animals are designed such that relative Resource Holding Power partially determines the resolution of resource conflicts

*Selection Pressure: more formidable organisms will expend their competitor's fitness budget faster than less formidable organisms, and thus will be more likely to gain contested resources.*

The effect of the selection pressures modeled in the AWA are evident in the animal behavior literature (for reviews see Archer 1988; Huntingford and Turner, 1987; Krebs and Davies, 1993). Animals of all species have conflicts over resources, and across a broad range of species less formidable organisms give way to more

formidable organisms. Formidability is usually indexed by the size of the organisms, but there are exceptions where weapon size is the best indicator of an organism's ability to inflict costs on a competitor. In these cases the organism with smaller weapons, even if it is larger in size, will relinquish the resource to its competitor (Barki, Harpaz & Karplus, 1997; Neil 1985). The advantage in resource conflict given to more formidable competitors has been found in species as distantly related as the sea anemone (Brace & Pavey, 1978), amphipods (Connell, 1963), beetles (Eberhard, 1979), African buffalo (Sinclair, 1977), crayfish (Hazlett, Rubenstein & Ritschoff, 1975), field crickets (Hofmann & Schildberger, 2001), green sunfish (Hale, 1956), mice (Ginsberg & Allee, 1942), golden hamsters (Marques & Valenstein, 1977), and New Forest ponies (Tyler, 1972). A particularly dramatic example was found by Petrie (1984) who studied territory size in the moorhen (*Gallinula chloropus*) and found that relative male weight was a perfect predictor (Spearman correlation coefficient of 1.0) of territory size. For many more examples see Archer (1988), chapter nine.

2. Animals are designed such that the relative valuation of the resource partially determines the resolution of resource conflicts

*Selection Pressure: Animals that value a contested resource more will have a larger fitness budget and thus be more likely to gain the contested resource.*

As the Asymmetric War of Attrition suggests, relative Resource Holding Power (RHP) is not the only factor that impacts resource conflict. If a resource is of differential value to the two organisms, the organism that values the resource more

will be more likely to receive the resource because this organism will be willing to incur more costs during the struggle for the resource. Because of the differential fitness budget, more formidable organisms will sometimes be selected to defer to organisms that place more value on the contested resource. The results of these selection pressures can be seen most easily in experiments in which food is withheld from two animals for different amounts of time, after which they are allowed to compete to gain access to a food item that is valuable to both but, because of differential starvation, more valuable to the animal that has had food withheld for a longer period of time. As predicted by the model, animals that have been deprived of food for longer periods of time tend to win such contests. This basic effect has been replicated with white rats (Bruce 1941), chimpanzees (Nowlis 1941), house cats (Cole & Shaffer 1966), hermit crabs (Hazlett 1966), cray fish (Hazlett et. al. 1975), bald eagles (Hansen 1986), and dark-eyed juncos (Cristol, 1992) among other species. The effect is not limited to differential valuation of food, but has been found for differential value of mating opportunities, nesting sites, feeding sites, and webs (for a review see Enquist & Leimar, 1987, for a particularly clear example see Austad, 1983).

### 3. Animals are designed to assess relative RHP and relative resource value

*Selection Pressure: Given that one animal will win a resource conflict and that there are costs incurred as a result of engaging in conflict, animals will be selected to assess cues that predict who is likely to win, and cease costly competition if they are unlikely to win.*

The data described above (see points 1 and 2) demonstrate, at a minimum, that a broad range of animals compute their own RHP (i.e.  $k(X)$ ) and the value they place on the resource (i.e.  $v(X)$ ) and use this information to govern their aggressive behavior in situations of resource conflict. A well-designed mechanism for adaptively navigating the selection pressures posited in the AWA would also estimate these parameters in the other contestant. Knowledge of the other animal's RHP and the value it places on the contested resource could be used to estimate the probability of winning the fight and be used to make prudent decisions regarding the instigation or continuation of a violent contest.

One could posit a number of different computational mechanisms designed by natural selection to compute another organism's RHP. For example, their strength could be gauged during physical confrontations by the amount of damage done to the assessor during the conflict per unit of time (e.g. if you're an ant and you've been fighting another ant for two seconds and have already lost two limbs, this is evidence that the opponent has relatively high fighting ability). Other possible designs for an RHP estimator include an estimate based on visually accessible phenotypic traits, analysis of vocal emissions for cues of body size, memory for past encounters with this particular opponent, deference given to the competitor by conspecifics of similar RHP, chemical cues of immunological health, etc. Given the ubiquitous nature of resource conflict and the intense cost of misperceived RHP in others (e.g. losing a resource one could have won or losing a violent interaction one could have avoided), it is likely that natural selection has designed extremely sophisticated machinery into

the nervous systems of animals that is designed to accurately assess the likelihood of victory in situations of resource conflict.

Evidence exists in a variety of species indicating that animals have not only been designed to calculate their competitor's RHP, but they are designed to broadcast their own as well. Several species of birds and lizards have evolved physical signals of fighting ability (i.e. "badges") that appear to track, quite accurately, the organisms' fighting ability (Rohwer & Rohwer, 1978; Evans & Hatchwell, 1992). These badges would be useless if the signals were not used by their competitors to compute their RHP. Furthermore, the widespread existence of dominance hierarchies in many species indicates the capacity of such animals to measure their RHP with reference to others (for a review of the non-primate animal literature see Huntingford and Turner, 1987, for primate examples see Smuts et al., 1987).

More direct experimental evidence of RHP estimation comes from experimental studies such as Richard Alexander's study of fighting crickets (1961). When crickets were allowed to physically dominate "dummy" crickets (i.e. artificial animals created by biologists to resemble crickets) the victors become more likely to challenge live competitors in the future, demonstrating that victory is tracked in the brain of the winning animal as an indicator of the likelihood of future success. The same results were found using a slightly different method with mice and rats (Scott & Fredericson, 1951).

Evidence that animals estimate relative resource value is less common. This lack of evidence stems from at least three factors.

First, in many cases, the only phenotypic manifestation of the value placed on a resource resides in the neural structure of the organism. For example, the value male dung flies place on a receptive female declines as a function of the length of time that the dung fly has spent inseminating the female (Parker, 1978). After approximately 40 minutes the benefits of further insemination are outweighed by the probability of finding a new female and the dung fly ceases copulation and begins searching for additional females. If one male dung fly discovers a competitor currently mating with a female, the information that would indicate the relative value of the female is not available to the new male because he does not have any indication of how long his competitor has been mating with the female. The information is represented in the neural structure of the mating male; but such information is difficult to perceive. It is likely for these reasons that RHP, often indicated by body and weapon size which is more easily perceived, is estimated more accurately and with greater frequency than relative valuation.

Second, resource valuation often correlates negatively with RHP. Starvation, for example, will lead to a higher valuation of contested food items but simultaneously to a diminution of fighting ability. Similar situations occur when weak males of polygynous species, who are unlikely to mate at all, face conflicts over access to a reproductive female. The weak male has low RHP but the value of the resource is greater as life-threatening injuries are of almost zero reproductive consequence. These negative correlations will obscure data in ways that make it difficult to detect mechanisms for computing relative resource value, but will not remove the selection

pressures themselves. In other words, animals that are better able to detect resource valuation will still be at a reproductive advantage compared to animals that lack such machinery because they will be able to make more prudent decisions when faced with competitors for whom the negative correlation between resource valuation and RHP does not hold.

Third, methodologically it is difficult to distinguish between an animal who relinquishes a resource because it perceives its opponent to value it more and an animal who relinquishes a resource because it perceives its opponent to be more formidable (i.e. higher in RHP). For example, if two beetles are fighting over some food and the stronger of the two has broken its competitor's carapace and severed one of its antennae, yet the competitor continues to fight, a well-designed beetle would be able to estimate that the value placed on the contested resource by the highly-injured beetle would be quite high. If it were of little value the beetle would have left before receiving such injuries. Injuries, particularly internal injuries and energy expense, are difficult for researchers to perceive, quantify and evaluate in most animals. Without quantifying them, all researchers can record is that animals with less RHP will win conflicts over resources if the ratio of valuation is large enough. This has been shown (see point 2), but it is not currently possible to determine if this is because the larger animal runs out of fitness budget first and leaves, or if the larger animal estimates (perhaps based on an assessment of its opponent's injuries) the relative value of the resource and leaves before the fitness budget is expended.

Those reasons notwithstanding, there is at least one well documented case of an animal being designed to estimate relative resource value and using that to resolve conflicts. Male elephants go through an energetically expensive cyclic fertile period called *musth*, during which a sexual mating will be much more likely to yield offspring (one can think of this as male ovulation). Joyce Poole (1998) has demonstrated that non-musth male elephants will relinquish a reproductive resource (i.e. a mateship) when faced with a male in musth even if the nonmusth male is somewhat larger (though not if the nonmusth elephant is much larger). This is not because the mateship is entirely without value to the non-musth elephant, as males not in musth will still mate with the female and will resolve competition for her based on relative formidability.

4. Animals are designed to demonstrate RHP to lower the costs of contests

*Selection Pressure: Both organisms engaged in a costly contest over a resource will minimize that cost if the eventual loser can recognize that it will lose.*

When the AWA equation is satisfied, organism Y would be selected to relinquish the resource. If both organisms have perfect information about the value of the resource to themselves and their opponent, and perfect information about relative RHP, there should be no such thing as animal aggression over resources. The organism that would lose the fight would simply relinquish the resource and neither organism would need to pay the cost of the contest. Most aggression takes place, according to G. Parker and other evolutionary biologists, because organisms do not have perfect information about contests (Parker, 1974; Enquist & Leimar, 1983;

Enquist et al. 1990). Much animal aggression, when it comes to contested resources, is not the result of adaptations designed to completely eliminate one's opponent but of adaptations designed to compute relative formidability in order to determine which organism will get the contested resource. When sufficient information is gathered for an organism to realize it will lose, that organism will retreat. Aggression can thus be characterized as information exchange in which organisms communicate and assess each other's ability to inflict costs on the other.

Animal aggression over resources that is designed to demonstrate relative RHP should follow a general pattern of escalation in which low cost demonstrations (which are probably less accurate) are exchanged before higher cost demonstrations or eventual non-ritualistic, "no holds barred" combat. Large discrepancies in RHP should be evident even by comparatively inaccurate demonstrations of RHP. For example, tail beating in the cichlid fish (i.e. shaking one's tail at the opponent so that it makes waves that hit them) is an indicator of body size and strength that is presumed to be 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 (Enquist et al., 1990), bowl and doily spider (Austad, 1983), African buffalo (Sinclair 1977), beetles (Eberhard 1979), common toads (Davies & Halliday 1978), red deer (Clutton-Brock & Albon 1979) and pigs (Jensen, P. & Yngvesson, 1998).

If a class of animal aggression is designed to demonstrate RHP, one would predict that such aggression, contrary to predator aggression for example, would be designed to end with signs that the competitor is relinquishing the resource. It is obvious that animals in these contests do not take victory as an opportunity to kill their competitors, but the evidence of submission as an end-point of animal aggression is more convincing than the absence of continued conflict. 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 relinquishes the resource. Such submission displays are common in a variety of species; for example, hens (Maier, 1964), cichlid fish (Jakobsson et al. 1979), and hamsters (Faruzzi et al. 2005), and they are found almost everywhere in the primate order (Smuts et al. 1987).

5. Animals are designed to demonstrate their valuation of the resource to lower the costs of contests.

*Selection Pressure: Organisms engaged in a costly contest over a resource will minimize that cost if the eventual loser can recognize that it will lose.*

For reasons mentioned earlier, it is difficult to gather evidence that animals are computing the relative valuation of resources even though there is reason to believe they do so. In the one documented case, elephants, in which animals clearly assess the value their competitor places on the resource, there is also evidence that the contestants signal that valuation. To recap, male elephants compute the relative value of a female based on the current musth-state of the competitive males. There is abundant evidence that natural selection has designed signaling mechanisms in the

elephant to demonstrate their valuation of the resource (i.e. their musth state). To quote Joyce Poole, “Musth is characterized by a distinct posture, swollen and secreting temporal glands, dribbling of strong-smelling urine, and a particular, very-low frequency, vocalization, the ‘musth rumble’” (1999). The redundancy built into the signal (note that the signal contains visual, auditory and olfactory components) is a testament to the intensity of the selection pressure to demonstrate the value of the contested female to the male.

6. Species in which animals typically engage in repeated interactions will maintain an internal representation of relative RHP that governs conflicts of interest without needing to re-establish RHP through competitive interactions.

*Selection Pressure: Repeated contests to demonstrate RHP are costly.*  
*Organisms that can store this information for future use can avoid the costs of repeated contests.*

Due to the cost of RHP demonstrations, there would be selection to store information about the RHP of competitors in species with repeated encounters (i.e. social species). These internal representations would then be consulted with the appearance of each new resource conflict so that the costs of reassessment would be removed. A species equipped with such a mechanism would have stable, long-term patterns of interaction in which animals would appear to spontaneously relinquish resources to certain individuals (those with higher RHP) while expecting other individuals (those with lower RHP) to relinquish resources to them. Such interactions are called dominance hierarchies, and have been demonstrated in many social species

such as hens (Maier, 1964), dark-eyed juncos (Cristol, 1992), Panamanian insects (i.e. *Zorotypus gurneyi*, Chloé, 1994), chickens (Guhl, 1956), dark chub fish (Katano, 1990), hyenas (Owens & Owens, 1996), cockroaches (Ewing, 1967), and every social monkey and ape (Smuts et al. 1987).

It is important to point out that these dominance hierarchies are often established in the absence of any immediate resource conflict. This may be a design feature that allows relative RHP to be assessed without confounds: if there were a contested resource, then ferocity of fighting would reflect not just RHP, but also the competitors' relative valuations of the resource. In the absence of any resource, the fight reflects only relative RHP. Indeed, conflicts designed to demonstrate RHP often take place *before* times of great resource conflict. For example, male elephant seals spontaneously begin conflicts of assessment about a month before females even arrive on the beach (LeBoeuf 1972, 1974). Pigs (Jensen, P. & Yngvesson 1998), elephants (Poole 1999) and chickens (Maier, 1964) spontaneously engage in conflicts of assessments whenever they lack information about another conspecific – even in the absence of a contested resource. Maier exposed caged chickens to each other in ways that prevented conflicts of assessment but enabled visual and auditory contact. Chickens were released after three weeks and spontaneously entered conflicts of assessment. It is not novelty in chickens that leads to assessment, but lack of information about RHP. Jensen and Yngvesson (1998) did the same experiment with pigs and found that conflicts of assessment took place, but that the early stages of conflict were shorter (presumably because the pigs gleaned some information about

each other from being within visual contact) but that later stages of the conflicts were just as long.

***C. Conclusion about animal aggression: Animals are designed to compete over resources by determining and demonstrating relative RHP and the relative valuation of the contested resource in ways that minimize the costs of costly competition.***

The AWA summarizes a number of selection pressures into a single framework that has made successful predictions about animal competition and conflict in scores of diverse species. It would be odd to assume, a priori, that these selection pressures did not have some impact on human adaptations for conflict and aggression. The research reported herein proceeds from the assumption that these selection pressures have shaped the design of mechanisms in the human mind that are designed to navigate the problems inherent in resource conflict.

***D. A cognitive model of animal aggression***

Traditional evolutionary biology does not rely heavily on cognitive models of behavior. At times, biologists simply connect adaptive behavior to selection pressures, without reference to internally represented variables that guide the adaptive behavior. It is important to make implicit assumptions about these internally represented variables explicit, in a way that can be tested. Otherwise, it is difficult to explain nonadaptive behaviors that are particularly common in species that no longer live in the environment they are adapted to, such as humans. It is worth taking a moment to

construct a rudimentary cognitive model, applicable to almost any social species with frequent conflicts of interest, based on the selection pressures in the AWA.

The AWA is a model of selection pressures, not computational mechanisms, and as such the value of a resource is denoted in terms of fitness differentials. A computational model - a model of the variables used by an organism over ontogenetic time - cannot include fitness unless the animal is assumed to have a perfectly accurate fitness assessment mechanism, of which no evidence (theoretical or empirical) exists. The computational machinery instantiated in animal brains for regulating conflicts of interest do not compute fitness exactly, but rather compute an index I will call **welfare**, which is designed to approximate fitness. This index, over evolutionary time, had to demonstrate a substantial correlation with fitness, but may or may not track fitness in current environments. For example, it is not an accident that humans consider the insertion of sharp objects into their eye to be a cost because such behavior has for some time been a detriment to fitness. But one should not expect (in humans or other animals) that all resources internally represented as valuable for the animal do in fact increase reproductive fitness. Males will fight for females who, unbeknownst to the contestants, are infertile. Dogs will fight over meat laced with strychnine. Reproductive fitness over evolutionary time must, for scientific purposes, be kept distinct from its internally represented proxy (i.e. welfare) if one is to understand the functioning of the organism. In *all* cases of organismic behavior, it is the neurally-represented mechanism that decides behavior (i.e. muscle movement) and not the selection pressure directly.  $V(X)$  is a selection pressure;  $Welfare(X)$  is a

cognitive representation that was statistically correlated with  $V(X)$  over evolutionary time.

Resource Holding Power is denoted in terms of relative fitness loss during contests over resources (i.e. who loses future copies of their genetic code faster during a contest). Organisms, of course, do not have perfect representations of that rate of loss, and thus must create and store some index of RHP. I will call this index **formidability**: an internally-represented variable that is designed to approximate an animal's own ability to inflict damage to others' welfare and withstand damage to its own welfare.

### *E. Humans: constructing a cognitive model of human anger*

The scientific power of comparative analyses is not derived from demonstrating similarities between humans and non-humans<sup>3</sup> but from distilling abstract but precise models of selection pressures that can then be used to make predictions about humans taking into account unique aspects of human evolutionary history (Tooby & DeVore, 1987). Humans, I propose, engage in many conflicts of interest—perhaps more than any other species—due to at least two unusual ecological circumstances that were part of human evolutionary history.

Humans are a highly social species demonstrating coordinated actions with large groups of individuals in their natural environment (i.e. hunter-gatherer ecology; see Kelly 1995). Repeated encounters with hundreds of individuals were not uncommon,

---

<sup>3</sup> “I didn't know chimps do that... maybe that's why humans do.” It would be just as reasonable to say “I didn't know humans do that...maybe that's why chimps do.” Similarities between species are not mutual explanations, but opportunities to derive common selection pressures that could account for the evolution and maintenance of the trait (Tooby & DeVore, 1987).

and most importantly, the encounters frequently involved coordinated action during which conflicts of interest would have been ubiquitous (e.g. who would do what, when, for how long, for what benefit, to be distributed how, to whom, under what circumstances, etc.) Humans currently living in environments similar to the environments that shaped human ancestors over the last 5-10 million years engage in coordinated action for almost all life tasks including: hunting, gathering, defense from predators and hostile humans, offensive warfare, food preparation, child care, exploration, shelter construction, large tool making (e.g. canoes, bridges, etc.), and enforcement of laws and rules (see Shostak 1981 for an easily readable account of hunter-gatherer life; see Kelly 1995 for a more comprehensive, quantified account of the various foraging lifestyles).

The limits of meta-cognition (i.e. the ability to represent and consider information that is not currently available to the senses, such as pretending, imagining, planning, predicting and considering “what would have been”) in other animals are still being explored, but it is generally accepted that humans have a far greater capacity to meta-represent than other animals, and it is certainly true that humans have a greater ability to communicate those hypothetical representations to others (see Cosmides & Tooby, 2000). This presents a vast landscape of conflicts of interest that are invisible to other animals. Dung flies can fight to prevent another from inseminating their mate, but humans can fight over someone planning to inseminate their mate, or over someone planning on going somewhere which would make a third person more likely to

---

inseminate their mate, or over someone planning on going somewhere which would make a third person more likely to tell something to a fourth person who would then be less willing to help fight to prevent a fifth person from inseminating a mate, and so on. Furthermore, humans could more easily fight over inaction – why someone didn't stop someone from stealing her meat.

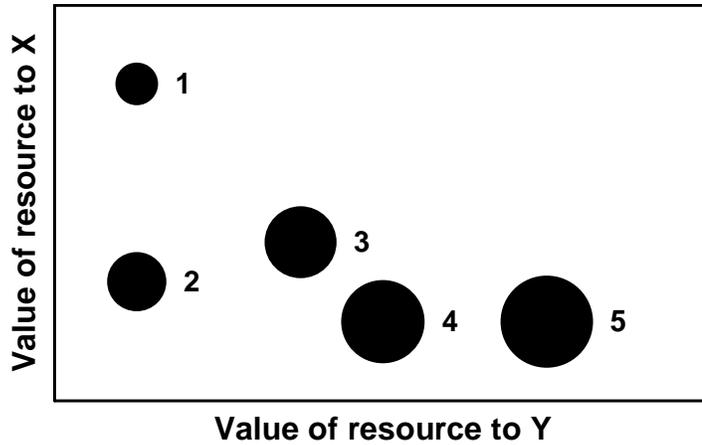
For these reasons, human resource competition often looks different from non-human animal resource competition. Though humans face conflicts of interest over morsels of food (e.g. “that’s the last doughnut!”), the majority of resource conflicts between humans do not involve tangible, materially-represented resources, but instead involve conflicts over courses of action (retrospective and prospective), exchanges of information, social alliances, and other abstract cost-benefit tradeoffs between the welfare of two individuals. Such tradeoffs fit the mathematical definition of resource competition; and thus would be subject to the same selection pressures as the hypothetical morsel of food in the traditional AWA equation.<sup>4</sup>

I will refer to resource competition as **cost-benefit transactions** when theorizing about resource competition with humans. Cost-benefit transactions can be classified according to the ratio of cost to benefit (i.e. relative valuation of the contested resource) that results from the imposition of the cost on agent X for the benefit of agent Y. See Figure 1-1.

---

<sup>4</sup> A resource in the AWA (e.g.  $v(X)$ ) is a differential fitness state between two outcomes: the fitness of X if they receive the resource minus the fitness of X if they do not receive the resource. In the same sense, a fight over whether a friend of mine should have proofread this paper can be represented as follows:  $v(X)$  = my fitness after having my paper proofread minus my fitness after not having my paper proofread;  $v(Y)$  = my friend’s fitness after not proofreading my paper minus my friend’s fitness after proofreading my paper.

**Figure 1-1. Examples of cost-benefit transactions as ratios of value.**



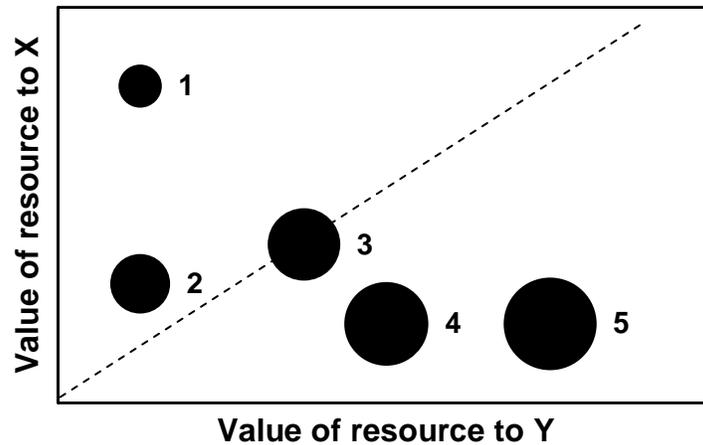
In Figure 1-1, there are five instances of cost-benefit transactions that differentially affect the welfare of agents Y and X. The size of the circle indicates the cost-benefit ratio (i.e. value (Y) / value (X); the extent to which Y values the resource more than X). Such examples could be: 1) Person Y borrows X's car and crashes it to avoid getting a \$20 speeding ticket (small benefit for Y, large cost to X); 2) An airline employee (Y) makes X wait for fifty minutes so that the captain can finish his birthday cake. 3) Y takes the last doughnut at a meeting between X and Y (equal value to both); 4) Y ruins X's favorite sweater by using it to wrap Y's dog's leg after it's caught in a bear trap; 5) A supermarket checkout employee (Y) makes X wait for 2 minutes while Y takes a phone call to discover how his parent fared in open-heart surgery (benefit to Y much higher than cost to X).

#### 1. Welfare Tradeoff Ratios (WTRs)

Humans use a Welfare Tradeoff Ratio (WTR) that indicates the extent to which they will allow costs on themselves in order for another to benefit and the extent to which they will impose costs on that other for their own benefit.

It is posited that humans internally represent a threshold for acceptable cost-benefit transactions called a Welfare Tradeoff Ratio (i.e. for agent X, the value  $v(X)/v(Y)$ ) indicates the cost-benefit transaction, below which Y will get the resource and above which X will get the resource). An individual will have different Welfare Tradeoff Ratios for different individuals. For example, a person might be willing to ruin a stranger's favorite sweater to dress their pet's wound, but unwilling to ruin the Dean's sweater for a similar reason. An individual may be unwilling to give several thousand dollars to pay for a stranger's education but willing to spend that amount on their child's education. This ratio can be seen as a threshold (the dotted line in Figure 1-2), above which agent Y would relinquish the benefit rather than impose the cost on X, and below which Y would impose the cost on X to receive the benefit. This ratio is Y's Welfare Tradeoff Ratio with respect to X – as it indicates the extent to which Y is willing to tradeoff his own welfare in order to raise X's. In this case, he would not be willing to ruin X's car to avoid paying a speeding ticket (1 – is above the line), but he would be willing to make X wait in line while he called the hospital to check on a loved one (5 – is below the line).

**Figure 1-2. Welfare Tradeoff Ratio as a threshold of cost-benefit transactions.**



From a selective point of view, and ignoring cases of kin and the complicated nature of human cooperation, natural selection can be expected to have designed humans to adopt the lowest possible Welfare Tradeoff Ratio toward others (i.e. impose any cost in order to get even a meager benefit) – just as non-human animals would be designed to gain any resource regardless of its value to the competitor. But given the costs of contests and the nature of kin and competition, natural selection is predicted to have designed humans such that Welfare Tradeoff Ratios will be set higher based on numerous factors related to another’s ability to enforce their own welfare. One set of factors is related to the ability to enforce WTRs by threatening to inflict harm: these include, e.g., greater physical strength and more coalitional support. Welfare Tradeoff Ratios set primarily by the threat of force will be consulted, presumably, only when there is some possibility that the individual will be present to defend his or her interests. The other set of factors is related to the ability to enforce

WTRs by threatening to withdraw the benefits of cooperation: these included, e.g., the person's status as a frequent and dependable reciprocation partner, their status as a friend who has a stake in your welfare (Tooby & Cosmides, 1996), or their possession of special abilities that can be deployed to benefit others. Another factor, independent of the ability to enforce higher welfare trade-off ratios, is kinship: the welfare of one's children and kin should be valued more highly than that of nonkin, in ways that satisfy Hamilton's rule (Hamilton, 1964).

It is likely that there are at least two types of Welfare Tradeoff Ratios that govern cost-benefit transactions in different contexts: 1) extrinsic Welfare Tradeoff Ratios – which define the threshold of cost-benefit transactions when both parties are present or otherwise capable of defending their interests, and 2) intrinsic Welfare Tradeoff Ratios – which define the threshold of cost-benefit transactions when the other individual is not present or unable to defend their interests. Intrinsic WTRs allowed individuals to adaptively partition cost-benefit transactions in a world where the welfare of other individuals was of adaptive significance for oneself.<sup>5</sup>

Welfare Tradeoff Ratios are hypothesized to be internal regulatory variables that govern choices that pit one's own welfare against someone else's. These variables are computed from a number of factors, including the other individual's genetic relatedness to oneself, their reproductive value, their value as a cooperative partner and associate, and their formidability relative to one's own. This last variable, the

---

<sup>5</sup> Though the full selection pressure analysis of intrinsic WTRs is beyond the scope of this paper, a quick starting point would be as follows: a subset of individuals in one's social world can improve one's welfare as a result of their existence, social power and well being (e.g. a devoted friend, a

formidability index, is presumed to be phylogenetically and functionally related to RHP-estimators that non-human animals use to determine resource division with conspecifics in dominance hierarchies. In these cases, organisms reference an internally represented variable that determines the kinds of resource divisions that are acceptable.

2. The Regulator Theory of Anger: Human anger is an adaptation to raise another's Welfare Tradeoff Ratio with respect to oneself

An individual human is born into a world populated by others who, under most circumstances, are designed to give as little as necessary to the individual (i.e. to have low Welfare Tradeoff Ratios [WTRs] toward others, such that they would be willing to impose large costs on that individual for even small benefits). There would have been constant and powerful selection for an individual to raise the WTRs that others used when interacting with him or her so that they would forgo some subset of cost-benefit interactions that imposed costs on the individual. The higher the other's WTR can be pushed up, the more costs one can avoid.

Thesis: It is proposed herein that a large and well bounded subset of phenomena that people refer to when they use the word "anger" can be understood as the output of a highly-sophisticated, complex, reliably-developing computational system, instantiated in neural tissue and designed by natural selection, that is deployed as a negotiative tool to resolve conflicts of interest in the present or future in the angered individual's favor. The anger system does this by initiating behaviors that (1)

---

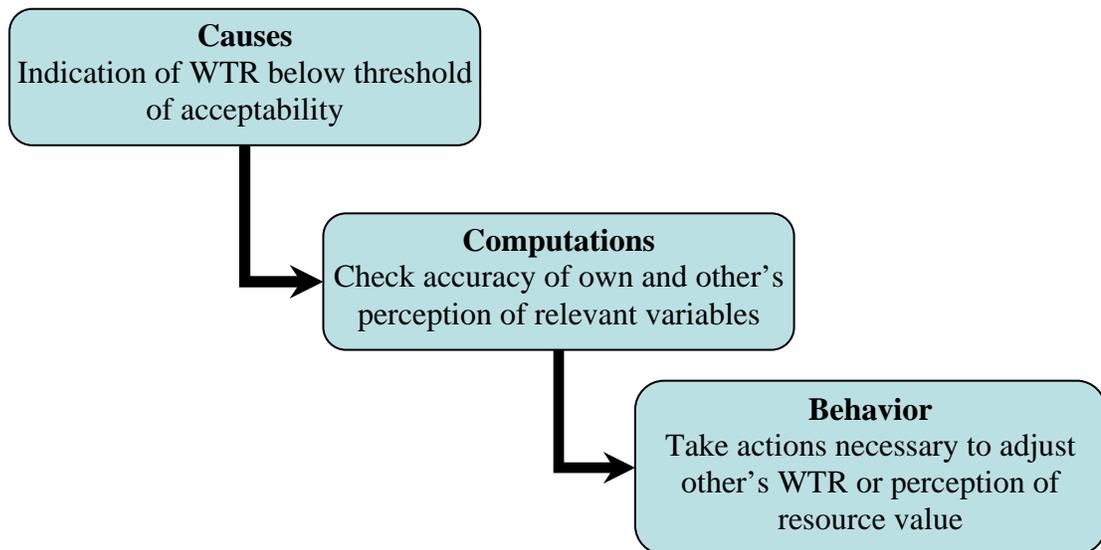
caretaker of one's children, a generous acquaintance, etc.) Benefits given to them, even without their

recalibrate the target's estimates of the costs and benefits of actions to the target and to the angered individual, and (2) raise the target's Welfare Tradeoff Ratio toward the angered individual, so the target takes that individual's welfare more into account in the present and/or future. Two main negotiative actions deployed to recalibrate the target's Welfare Tradeoff Ratio are (1) threatening to inflict costs (or actually doing so) and (2) threatening to withdraw cooperation (or actually doing so).

---

knowledge, will correspond to benefits to oneself.

**Figure 1-3. The Basic Computational Process of Anger**



### Causes of Human Anger

Anger should be triggered when one individual interprets the actions of the target as indicating that the target's WTR toward the individual is lower than the accepted level. It may also be triggered when one's ability to inflict costs or confer benefits has increased, in an effort to negotiate a new equilibrium value.

In this view, an important cause of anger should be an unambiguous display of a low WTR with respect to the individual who becomes angry. This can be indicated in a number of different ways (some of which will be discussed in Chapter 2), but the most theoretically clear indication of a low WTR is the target's willingness to take actions that impose a large cost on the angry individual in order for the target to receive a small benefit. Holding other variables constant, anger will be more likely to be activated over a cost-benefit transaction as i) the cost imposed on the individual

increases, ii) the benefit reaped as a result of that cost decreases, and iii) characteristics of the instigator, the angry individual and their relationship indicate that it is possible for the angry individual to force the other to use a higher WTR than was indicated by the cost-benefit transaction imposed.

#### Computational processes executed by Anger

Just as animals are designed to minimize the cost of violent confrontations that can be avoided by communication regarding relative resource value and relative RHP, so should human anger be designed to minimize the costs of anger that are triggered by misperceived costs and benefits on behalf of the angry individual. For example, imagine Joey became angry when he was made to wait half an hour for a colleague, Bob, to pick him up at the airport, because Bob was talking to his mother on the phone. At first it appears that Bob is indicating a low WTR with respect to Joey; he is imposing a moderate cost (i.e. half hour delay) for a small benefit (i.e. chatting with his mother). Were Joey to immediately conclude that Bob's actions reflect a low WTR and proceed to attempt to recalibrate Bob's WTR by threatening to inflict harm or withdraw cooperation, future cooperative effort (such as mutual airport pickups) would be placed in unnecessary jeopardy – especially if Bob had been talking his mother out of suicide (i.e. Joey misperceived the benefit to Bob of continuing to talk with his mother as small, when in fact it was quite large).

When evidence clearly exists that the instigator perceived the cost to be large and the benefit to be small and enacted the transactions with such knowledge (i.e. intentionally; see Chapter 2), then anger should trigger behavior that is designed to

increase the WTR in the target such that they will be less likely to impose cost-benefit transactions of that ratio (or worse) again.

On the other hand, when evidence exists that the target perceived 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 a subset of other transactions involving the same misconstrued variable) from being enacted. For example, imagine Joey left Bob's front door open while having a business lunch in order to get a breeze, and Bob's cat ran away and was found two days later. Joey thought the cost to Joey of leaving the door open was very little, perhaps nothing, but the cost was actually high. To the extent that Bob can make Joey recognize the actual cost, Joey will be less likely to prop the door open next week while moving in a cabinet, or to keep the door open while saying goodbye to friends, etc.

#### Behavioral Responses generated by Anger

The primary function of anger is to raise the magnitude of the WTR of an individual who has demonstrated a lower WTR than is acceptable to the angry individual, and/or to recalibrate that individual's estimates of the magnitude of the costs imposed and benefits received. Internal regulatory variables in the target should be open to modification when this will allow the target to avoid being harmed or having cooperation withdrawn—more precisely, in circumstances that predicted, ancestrally, that these two negative outcomes were likely. In other words, the target's

WTR should change in response to credible threats of harm or withdrawal of cooperation.

Thus, when the anger system is triggered by evidence that the target's WTR toward oneself is too low, it should motivate the angered individual to make credible threats, or demonstrate qualities that would make such threats credible, if made. WTRs are hypothesized to be set, partly, by relative formidability, just as it is with many other animals. The extent to which an individual can damage you is predicted to affect the kinds of cost-benefit transactions you will impose on them; and thus if an individual is showing evidence of a low WTR, it could be the result of an underestimation of one's formidability that could be recalibrated by a demonstration of said formidability.

As with non-human animals, formidability should be demonstrated starting with low-cost, presumably less accurate, demonstrations of physical strength and escalate as needed to more accurate and dangerous demonstrations of strength. Evidence reviewed in Chapter 2 reveals that this is the case.

The theory also predicts that anger should be designed to manipulate the other individual's estimates of the magnitude of costs and benefits inherent in the transaction. To the extent that you can increase another's perception of a cost he or she imposed on you, you can decrease the probability that the individual will impose such a cost on you again (e.g. "I liked that sweater a lot.") The same is true of reducing another's perception of the benefit they received (e.g. "you could have used

the paper towels on top of the refrigerator.”) Experiments in Chapters 3 and 5 test these hypotheses.

### 3. Summary of Regulator Theory of Human Anger

Selection pressures modeled in the Asymmetric War of Attrition have created mechanisms in nonhuman animals that use the relative welfare change and the relative formidability of the contestants to determine resource division. These same selection pressures are theorized to have played a substantial role in the design of human anger, resulting in computational machinery that functions to modify the social world in ways favorable for the angry individual.

Humans are theorized to have a set of internal regulatory variables, Welfare Tradeoff Ratios, which index the extent to which one is willing to impose costs on a given interactant for one's own benefit. They are also hypothesized to represent the WTRs others have toward themselves, which index the extent to which one can expect those others to impose costs on oneself for their benefit. Anger is theorized to increase the Welfare Tradeoff Ratios of other individuals who give evidence of having WTRs toward self that are lower than expected, and to regulate others' perceptions of cost-benefit transactions.

The primary cause of anger is hypothesized to be a demonstration that indicates another has a low WTR, in other words, another is willing to impose large costs on you for relatively small benefits. Anger is theorized to modify others' perceptions of cost-benefit transactions and variables used to set WTRs such that the offending

individual is less likely to impose cost-benefit transactions of that magnitude in the future.

## Chapter 2.

### **Alternative theories and how they explain known features of anger.**

Given the practical value of information pertaining to anger, aggression and conflict, there has been a great deal of empirical data gathered on this topic over the last seventy years in the field of psychology. Though these studies have resisted incorporation into an overarching theoretical framework, they provide a rich dataset which can be used to test new theories of anger and aggression. This chapter will review some results from the existing anger and aggression literature and then review dominant theories of anger in brief detail so that predictions from them can be derived and tested in the studies described in Chapters 3, 4 and 5. This review will show that many of the stronger and more reliable effects found in the literature can be parsimoniously explained by the Regulator Theory of Human Anger.

The format of this chapter is as follows:

1. Known features of Anger
2. Alternative Theories of Anger

#### ***A. Known Features of Anger***

Given the breadth of data collected on human anger, the first step when proposing a new theory must be to determine its consistency with empirical findings that have been shown to be both large in effect and robust across studies. Some of the basic features of anger that meet those criteria are reviewed below. It should be noted that while I did my best to choose the datasets below based on their effect sizes and replicability, it is not a complete review of the anger literature—left out are effects

that are small and/or unreliable. Theories of anger might be judged differently in light of alternative datasets that included these other effects.

1. Feature #1: Anger frequently results from the imposition of costs

Individuals tend to get angry when costs are imposed on them. This obvious feature of anger has been demonstrated across a range of costs. Most importantly, the magnitude of the anger response is proportional to the magnitude of the cost – imposition of a \$10 cost causes slight anger, while a \$10,000 cost causes much anger. Empirical studies that varied the magnitude of the cost have confirmed this across a host of different cost-types, such as the voltage of electric shocks (O’Leary & Dengerink, 1973; Helm, Bonoma & Tedeschi, 1972; Epstein & Taylor, 1967; Shortell, Epstein & Taylor, 1970) severity of insults (Taylor, 1967), crimes (based on sociological survey data on attitudes about crime; Blumstein & Cohen, 1980; Hamilton & Rytina, 1980; Warr, Gibbs, & Erickson, 1982; White, 1975), and monetary payoffs in economic games (Fehr & Gaechter, 2000). As I’ll explain below, this literature is crucial for evaluating models of aggression that posit negative affect as the primary (and sometimes only) cause of anger - such as the Frustration-Aggression Model, Berkowitz’s Cognitive Neoassociationism or Equity Theory.

2. Feature #2: Intentionality increases the likelihood of anger

A number of theoretical perspectives have posited intentional harms as the cause of anger and not simply negative affect, and these predictions have survived a number of empirical studies (see Epstein & Taylor, 1967; Ohbuchi and Kambara, 1985; and Nickel, 1974). When a cost is imposed, but without prior knowledge on behalf of the

imposer, there is much less anger and retaliation than if the cost was knowingly imposed.

Intentionality has so far been taken as a given feature of the social world and not something in need of explanation. Theories that posit intentionality have used different definitions. For example Heider (1958) defined intention as a plan that guides action, Kaufmann (1970) defined an aggressive action as one which is known by the actor to have a non-zero chance of inflicting harm on the target, and Tedeschi and Felson (1994) defined an intentional action as, “an act performed with the expectation that it will produce a proximate outcome of value to the actor.” (pg. 164). Intentionality, being an unobservable feature of human social life, must be a category the human mind uses to classify types of actions, and thus must be discovered and explored rather than defined as a given feature of the world. When viewed in this light, the concept becomes more complicated than “a plan for action” or “an act that will produce an outcome of value.” A fully described theory of intentionality will have to answer, at a minimum, what information would have had to have been known for something to qualify as intentional. For example, if someone plans hitting you with a toy ball and expects that you will enjoy this as part of a game, but when you are hit you become very upset because you didn’t want to play, would the person’s act have been intentional? Heider’s definition cannot answer this; Kaufmann’s definition would imply that it is not intentional; Tedeschi and Felson’s would say it was.

### 3. Feature #3: Apologies mitigate anger

The most reliable way to reduce anger yet found in the empirical literature is to apologize (Frantz & Benningson 2005; Goffman 1971; Harrell, 1980; Harrell & Hartnagel 1979; Ohbuchi, Kameda & Agarie 1989; Riordan, Marlin & Kellogg 1983; Schlenker 1980; Schwartz, Kane, Joseph & Tedeschi 1978; Zechmeister, et al. 2004). The content of apologies varies, and a great deal of empirical work remains to be done on distinguishing “real” from “false” apologies, and why angry individuals are so sensitive to the difference (see Holtgraves, 1989).

4. Feature #4: Anger and aggression are often used by males to restore “Face”

Violent and homicidal aggression are everywhere and across all known time periods most common among young men (Daly & Wilson 1988). These fights are, by all accounts, largely the result of insults and attempts to save “face” or attain status by fighting other young males. This account of violent aggression among males has been shown, to some extent independently, by criminologists (Luckenbill 1977; Wolfgang 1958), sociologists (Williams 1980), social psychologists (Berg and Fox, 1947; Felson 1982; Nisbett & Cohen 1996) and evolutionary psychologists (Daly & Wilson 1988). Most impressively, a host of cultural anthropologists have documented the positive relationship between fighting ability and status in non-police societies, including the Yanamamo of Venezuela (Chagnon 1983), the Achuar of Ecuador (Patton unpublished), the Dani of Highland New Guinea (Sargent 1974), Mexican gangs (Lewis, 1961), the Montenegrins of Eastern Europe (Boehm 1984), Inuit/Eskimos (Balicki 1970), the Jivaro Indians of the western Amazon (Karsten

1935) and American gangs (Toch 1969). In each of those societies, threats to one's "face" or "status" were often the beginnings of a violent episode between young men.

5. Feature #5: Personal insults are one of the most reliable causes of anger

Though not usually the object of study, personal insults have been used in aggression research for forty years and have (in most cases) been shown to be sufficient causes of anger (Geen, 1968; Worchel, Arnold & Harrison, 1978). More ecologically valid examples of aggression have found that personal "insults" almost always precede homicides (Berg and Fox, 1947; Toch, 1969; Luckenbill, 1977; Daly & Wilson, 1988) and assaults (Felson, 1982; Felson, Baccaglioni, & Gmelch 1986).

6. Feature #6: Anger has cross-culturally universal features and neuro-physiological locality

Many features of anger have already been demonstrated to be universal in design, including facial expressions, physiological changes, behavioral responses, typical antecedents, and neurological reactions.

The anger expression has been found to be universal across individuals and cultures, including cultures with no historical Western contact (Ekman 1973, Brown 1991). It has been demonstrated in six-month old infants (Stenberg, Campos & Emde, 1983; Stenberg & Campos, 1990) and in congenitally blind children (Galati et al. 2003).

Physiological changes as a result of anger have been found to be similar across cultures tested (Rime & Giovannini 1986): Subjects across a broad European sample reported that anger felt unpleasant, warm, and frequently led to muscular tension.

These patterns of responses were dissociable from other emotions in all cultures. Ekman, Levenson & Friesen (1983) demonstrated that anger results in a unique pattern of physiological arousal that is distinct from other emotions; this pattern includes increased blood to the hands, increased heart rate and galvanic skin response. As far back as 1929 (Cannon) it was known that anger caused a general increase in blood flow to the muscles.

Behavioral responses that result from anger (other than the aggression data described above) have been shown to be similar across European countries. Specifically, anger often led to vocal changes involving increased volume and sometimes trembling, changes in movement quality, clenched fists and increased hand movement (Shaver et al. 1987; Wallbot, Ricci-Bitti & Baenninger-Huber 1986). Furthermore, Klaus Scherer and colleagues have categorized vocal transmissions of emotion and showed that when angry, fundamental frequency (roughly, pitch) often increases in mean and variability (Banse & Scherer 1996; Johnstone & Scherer 2000). This pattern is also distinct from other emotions.<sup>6</sup>

The antecedents of anger have also been shown to be cross-culturally similar. Recall the cross-cultural data on aggression being driven by insults to one's "face" reviewed in feature #4. Wallbott and Scherer (1986) further demonstrated similar antecedents of anger for a broad range of European countries.

---

<sup>6</sup> There is some evidence that anger vocalizations come in two forms: one with increased F0 range and mean and another without. It seems likely that, given the cost of inadvertently signaling aggressive intent, natural selection has designed different displays: one for anger with clear aggressive intent, and another for anger that signals the desire to argue but not to demonstrate physical formidability. To my knowledge, this remains untested.

Finally, there have been numerous studies isolating parts of the brain that are differentially activated by anger, demonstrating that its neural-circuitry develops in similar ways across individuals and across species and is distinct from other emotions. Anger has been known to rely on the amygdala since at least the 1960s (see King, 1961). More recently Panksepp (2000) theorized, based on brain imaging and lesion studies, that anger relies mostly on the medial amygdala, BNST (bed nucleus of stria terminalis), and the medial and perifornical hypothalamus. The anger/rage system is moderated primarily by acetylcholine and glutamate in ways that connect the amygdala and PAG (periaqueductal gray) with the hypothalamus (see Siegel & Schubert 1995). Finally a recent meta-analysis has also confirmed that testosterone tracks individual differences in tendencies toward anger and aggression with a correlation of approximately .20 (Archer, 1998; see also Dabbs, 2000).

***B. Alternative theories and their ability to account for the basic features of anger***

A comprehensive review of all models of anger is beyond the scope of this text, so a subset of theories were included for comparison based on their popularity in the literature, their ability to generate clear testable hypotheses and their ability to account for some basic features of anger.

Given that the effects reviewed above are highly intuitive and have been empirically documented over long periods of time, anger researchers for many years have been aware of these facts. This introduces a problem because predictions can sometimes be wrenched from theories retrospectively to account for data that is

problematic. To some extent this is necessary for scientific progress, as theories should be revised as additional information is gained. But one must distinguish between datasets that are clearly predicted by the theory, datasets that can be retrospectively fitted to the theory (and are consistent with the theory), and datasets that (sometimes contrary to the author's claims about a theory) clearly contradict the logic of the theory itself. Therefore, while reviewing the alternative theories below I have been careful to reason about the theory itself and not grant predictions to a theory merely because an author claims them for their theory.

Anger and aggression have been thoroughly studied in diverse scientific disciplines including social psychology, emotional psychology, criminology, sociology, cultural anthropology, economics, and health psychology. None of the theories advanced in these areas of science have been explicitly made consistent with the modern theory of evolution. Thus some do not speak to function, and others use concepts of function that do not map onto the evolutionary biologist's concept of adaptive function. Therefore, few theories attempt to explain the causes, computations *and* effects of anger, but instead focus on one of these features (usually the cause). When possible, however, all theories will be reviewed with respect to their predictions about cause, computation and effects of anger.

#### 1. Cognitive-Motivational-Relational theory (Richard Lazarus)

##### Overview of Theory

In contrast to drive reduction theories of anger (Freud 1963, Lorenz, 1966), Magna Arnold (1960) posited appraisal theories of emotion, which suggested that

emotions are responses to appraisals individuals make on their environment (i.e. emotions are the result of computations on perceptual stimuli). Richard Lazarus elaborated on this perspective with his Cognitive-Motivational-Relational theory of emotions, which suggests that anger (and other emotions) result from certain primary appraisals involving goal relevance, goal incongruence, and type of ego-involvement (Lazarus 1991). Though Lazarus's theory is not explicitly functional, it does detail the causes of anger, certain computations, and the behavioral responses, allowing one to reason about how the behavioral responses correct the problems posited to be the cause of anger.

According to Lazarus, anger results when five appraisals are made. A relevant (1) goal must be made incongruent (2) in ways that threaten the preservation or enhancement of self-esteem (3) as the result of blameworthy actions made by an external source (4) that is best corrected by attack (5). A sixth appraisal determines the expected future costs of aggression and increases the tendency toward anger if expected future costs are low. These six appraisals are shortened into what Lazarus calls the "core-relational theme" of anger: "a demeaning offense against me and mine." He goes on to say, "What makes us angry is that we have been taken for less than we want to be by someone who is being either inconsiderate or malevolent" (Lazarus, 1994, pg. 212).

These appraisals require a number of computational components to be analyzed including blame, imputed control, the likelihood that aggression will resolve the dispute, and the likely outcome if aggression is used. The effect of these

computations, according to Lazarus is as follows: blame increases the likelihood of anger, imputed control on behalf of the offender increases the likelihood of aggression, the increased efficacy of using aggression will increase anger, and to the extent that aggression is likely to lead to a positive outcome anger will be increased.

Finally, the primary behavioral outcome of anger is posited to be aggression. “Although it is often inhibited for personal and social reasons, few would argue with the proposition that the innate action tendency in anger is *attack* on the agent held to be blameworthy for the offense,” (Lazarus 1991, pg. 226 – italics in original).

There are a number of theoretical problems with Lazarus’s proposal that make it problematic as a computational theory of anger.

Lazarus describes some of the variables most relevant to determining if anger is expressed or not: blameworthiness, the imputed control, and the fact that individuals should be less angry at individuals who are capable of harming them (he places this in appraisal six, which determines the likely outcome of an attack). But because the theory does not specify a problem that anger is designed to solve, it does not provide a logical or coherent structure relating these variables to one another; instead, it relies on the fact that we, as humans, know their effect on anger. A laundry list of causes and effects and moderating variables is not a substitute for a coherent theoretical framework that can organize those effects and variables under one or a few functions (e.g. the mapping of every artery, vein, and chamber of the heart and a list of all physiological circumstances that modify heart rate does not negate the importance of understanding that the circulatory system is designed to distribute oxygenated blood

to the body). Such a list would be particularly inefficient in discovering new features of the structure under study because lists of cause and effects don't make predictions about variables not yet identified or studied (e.g. a mapping of the circulatory system doesn't predict that heart rate should increase while exercising, but the theory that the heart functionally oxygenates the body would). I refer to this as the "laundry list problem."

To use a specific example, Lazarus posits that individuals are not likely to become angry when the person who enacted the goal incongruence (offense) could not have acted differently. Intuitively, humans know this to be true, and there is good empirical evidence, reviewed later, that demonstrates the important role of intentionality. But there is no theoretical reason why this should stem from a theory about rectifying goal incongruence. Lazarus states it, but it does not follow logically from a theory.

The second problem with the Cognitive-Motivational-Relational theory stems from a phenomenon labeled "instinct blindness," a tendency for humans to not recognize the complicated computational work done by unconscious mechanisms (Cosmides & Tooby 2000). Humans find it easy to intuit the meaning of Lazarus's words, but as cognitive scientists it is hard to articulate in computational terms the meanings of words such as "demeaning," "offense," "taken for less," "inconsiderate, and "blameworthy." For instance, would it be demeaning for a roommate to wash his bathtub with his friend's favorite sweater? If so, why? What variables is the human mind plucking from that situation, and what calculations are being done on them?

Would it be demeaning if it were a ragged old bath-towel? Would it be demeaning if he were blind and thought it was his own sweater? Would it be demeaning if he had ruined the sweater dressing a head wound? Would it be demeaning if he thought it was your brother's? Intuitively we know the answer to these questions, but a theory of the causes of anger that rests solely on intuitive concepts shifts the burden of a computational map to those who study the intuitive concept.<sup>7</sup>

Finally, some of the variables that Lazarus appeals to are not possible to put into computational terms. For example, he suggests that someone is blameworthy if they were "capable of having acted otherwise." Invoking unbounded hypotheticals as a criterion eliminates one's ability to classify the behavior. Anyone "could have" done anything if some unspecified features of the universe were modified. His example of someone who could not have done anything different was a clerk who made her customer wait while she took a phone call informing her of an accident her son had been in. He argues that the clerk could not have acted differently, thus the customer would not be angry. Even limiting the hypothetical to changes in the woman's behavior, she still could have acted differently. She could have slammed down the phone and done her job in agony. In Chapter 3 I will attempt to explain why the clerk's behavior does not elicit anger in a more computationally tangible way.

#### Account of Features of Anger

---

<sup>7</sup> Of course all cognitive theories must posit variables that are not completely computationally defined by the theory itself (the current theory does not explicitly detail the means by which individuals determine the magnitude of a cost or benefit, for example). But there must be some computational work done by the theory other than labeling behavior that leads to anger as "blameworthy" and all behavior that does not as "not blameworthy."

The Cognitive-Motivational-Relational Theory clearly predicts that imposed costs should cause anger to the extent that they are “demeaning” to either oneself or to someone one cares about, and that intentional costs should be more anger provoking because the target is seen as more “blame-worthy” when they intended the action. The fact that insults reliably provoke anger is clearly predicted by this theory, provided that those insults are threats to one’s esteem; and for similar reasons this theory predicts that violent aggression between males across cultures should also result from conflicts over status and “face,” though it is not clear why physically beating another man (or killing him in extreme cases) restores one’s face or sense of self; nor why these types of aggression are not practiced among women.

The universality of many features of anger and the effect of apologies are consistent with Cognitive-Motivational-Relational Theory, which does posit that cross-cultural similar functions for each emotion.

In general, the Cognitive-Motivational-Relation theory, being functional in nature, accounts well for the major features of anger, but suffers from a lack of computational specificity.

## 2. Cognitive Neoassociationism (Berkowitz)

### Overview of Theory

Berkowitz extended the earlier Frustration-Aggression Model into a model that attempts to explain the causes of anger and aggression in terms of “negative affect” (Berkowitz, 1989, 1990). Specifically, “associative networks link specific types of feelings with particular thoughts and memories and also with certain kinds of

expressive-motor and physiological reactions.” Furthermore, “the activation of any one of the components in the network tends to activate the other parts as well” (Berkowitz, 1990, pg. 496). This analysis suggests that the causes of anger will be found in: feelings, thought, memories, expressions, muscles movements and physiological reactions. It is not clear why negative affect is the most common cause of anger as opposed to anger-linked thoughts, memories, expressions or muscle movements, but according to Berkowitz the model does predict this. Negative affect is hypothesized to activate two separate networks of reactions – one being feelings, memories, ideas and bodily changes associated with escape, and the other being feelings, memories, ideas and bodily changes associated with aggression. The relative strength of these reactions are said to stem from, “a variety of factors – genetic, learned, and situational” (pg. 496). After these initial reactions the individual then makes, “appraisals and causal attributions and considers what feelings and actions are appropriate under the particular circumstances” (pg. 497).

Berkowitz’s model is supported by a number of experiments showing small anger-inducing effects of variables that are not, according to most theories, predicted to cause anger or aggression. For example, the mere presence of a gun in a room leads to a more aggressive response against a confederate (Berkowitz, 1968), and making an angry face will often result in self-perceived feelings of aggression as well as physiological changes consistent with anger (e.g., Adelman & Zajonc, 1989; Ekman & Oster, 1979; Ekman, Levenson & Friesen, 1983). Physical exercise (Zillmann, 1979), sexual arousal (Baron & Bell, 1977; Yoram, Malamuth, Feingold &

Feshbach, 1974; Donnerstein & Barrett, 1978) and crowding (Freedman, 1975) have all been shown to increase the anger response in one respect or another.

Unfortunately a number of these empirical findings have failed to replicate or been contradicted by later studies (see Tedeschi & Felson, 1994 for a review), and so they must be treated with a degree of skepticism. Outside of this database, however, lie the more basic features of anger, such as the importance of intentionality or insults, which cannot be addressed by the model in a specific computational way. In fact, the whole of individual differences in propensity toward anger and fear have been subsumed under “genetic, learned, and situational”—a description that could perhaps be applied to most phenotypic features, not just anger. The behavioral responses of anger were similarly moderated by appraisals, attributions, considerations and “what feelings and actions are appropriate under the particular circumstance.” Again, the specification of relevant variables remains rather vague. If one cannot specify basic theoretical variables, then one cannot really test the theory.

#### Account of Features of Anger

Negative affect is presumed to result from larger costs, and thus Cognitive Neoassociationism clearly accounts for the data showing that larger costs provoke more anger. This data, however, confounds actual costs imposed with the intention to impose costs – and thus can support either negative affect models (such as Cognitive Neoassociationism) or intentional models (such as the Cognitive-Motivational-Relational theory) depending on what happens when costs imposed on an individual are dissociated from the intentions of the individuals imposing those costs. Such

experiments have been done and have shown that intention to harm rather than the harm itself best predicts retaliation. For example, Greenwell and Dengerink had a confederate administer electric shocks to subjects in a competitive game (1973; see also Geen, 1968). The voltage was varied between conditions as was the signal (supposedly indicating the voltage). Subjects' aggressive responses were predicted by the signal of voltage but not the actual voltage administered to them. This caused Greenwell and Dengerink to conclude, "While attack is an important instigator of aggressive behavior, it appears that the physical discomfort experienced by a person may be subordinate to the symbolic elements that are incorporated in that attack [p. 70]." They may have understated their conclusion, as there was *no* significant increase in aggression as voltage of the shock increased. These results are problematic for theories that rely on negative affect as the primary cause of anger. Instead of there being a straightforward relationship between size of cost and degree of negative affect, it forces proponents of such theories to circularly argue that intentional harms produce more negative affect than unintentional ones, with no theory as to why that might be (Berkowitz 1989). This may be true, but if the causes of negative affect are so loosely defined it ceases to be an explanatory construct and becomes a vacuous label that could, in theory, account for any possible cause of anger. In short, Cognitive Neoassociationism cannot account for the effect of intentionality on anger, or why varying the magnitude of a cost has little or no effect once intentionality is controlled for.

The effect of apologies is crippling to theories that posit anger as a functional response to negative affect; Cognitive Neoassociationism has a difficult time accounting for why verbal indications of anything reduce anger so dramatically. It is similarly hard to explain why individuals respond to insults with anger – though this could, retrospectively, be explained as provoking negative affect.

The universality of certain features of anger is predicted by Cognitive Neoassociationism, as the anger/fear system is theorized to be phylogenetically related to similar systems in other animals – though why those features are present in humans and other animals is unexplained.

### 3. Equity Theory

#### Overview of Theory

Equity theory was proposed as a general theory of human behavior that could be applied to many domains of human life (Walster, Berscheid & Walster, 1973; Berkowitz & Walster, 1976). Edward Donnerstein and Elaine Hatfield (1982) tested the fit between data on anger and aggression and the predictions derived from Equity Theory with mixed results. Since that time only two empirical studies (to my knowledge) have been published that directly test predictions of equity theory on anger and aggression (Hammock, et al., 1989 and Craig et al., 1993). Despite its smaller role in traditional aggression writing (equity theory received only one small paragraph in the most recent Handbook of Social Psychology's chapter on aggression, Geen, 1998, pg. 328), Equity Theory is computationally specific with respect to a great number of variables that are known to cause anger (such as cost-benefit

interactions, indications that another deserves more than you, the notion of intentionality, etc.) Furthermore, given that restoration of equity can be conceptualized functionally, the theory can make predictions about the causes and consequences of anger, individual differences in proneness to anger, and conditions under which anger will be assuaged.

Equity theory, as applied to anger, predicts that anger is designed, to some extent, to make outcomes more equitable. This makes a number of predictions about anger that are, at first glance, very likely to be true. One, anger ought to be caused by costs that are imposed on an individual. Two, anger should be more likely when that cost is only applied to the angry individual but not others. Three, the punishment desired by the angry person should be greater as the cost imposed on the angry person increases (feature #1). Four, being given compensation for the lost resource will mitigate anger.

One reason equity theory may have been all but abandoned by anger researchers is its inability to explain some of the exacerbating factors that have been shown to increase anger in laboratory settings (for example the effects of exercise, sexual arousal, crowding, violent primes, etc. – see Cognitive Neoassociationism above). Donnerstein and Hatfield were the first to point out that equity theory cannot account for such findings (1982). To dismiss equity theory for these reasons seems premature, however. No theory of anger can yet account for all the data on human anger. To test among the remaining theories requires an empirical triage in which the most important effects (i.e. the variables with the largest effect sizes) are singled out for

explanation before attempting to explain the smaller, less replicable effects. For example, that insults can trigger anger is a major, replicable phenomenon.

Perspective is lost when a theory of anger is required to explain why vigorous exercise may, in some studies, slightly exacerbates one's aggression in response to an insult, while the alternative theory, which accounts for the effect of exercise, cannot explain why insults have any effect on aggression to begin with.

#### Account of Features of Anger

Equity theory clearly predicts that larger costs should provoke more anger because they create a more inequitable distribution. It is unclear whether intentional acts that result in inequitable distributions should be more anger provoking than unintentional acts, as both result in the same magnitude of inequity, but the later is, presumably, more likely to indicate continued inequitable treatment, and Equity theory suggests that anger should be result from a "inequitable relationship" (Donnerstein and Hatfield, 1982, pg. 310). Apologies, though doing nothing to relieve the inequity that caused anger, could be retrospectively explained as attempts to indicate that the act that caused the inequity was an isolated incident that does not indicate an inequitable relationship.

Equity theory has a difficult time accounting for why insults (in the absence of any resource division) should cause anger – as there is no inequity. Similarly, the male-male aggression that stems over conflicts of "face" and "status" across cultures is hard to account for without positing additional machinery that computes the likelihood of future inequitable treatment. It fails to explain why these conflicts are

often between men rather than women and why they often happen between strangers (who one has no particular relationship, let alone an inequitable one).

Equity theory makes no predictions regarding the universal nature of aspects of anger other than that they will be caused by inequity.

#### 4. The Frustration-Aggression Model

##### Overview of Theory

The Frustration-Aggression model was one of the first computational accounts of human anger (Dollard et al., 1939). Its two central claims are 1) that all aggression stems from a kind of frustration, and 2) that all frustration leads to some kind of aggression. Frustration is couched in a mix of behaviorist and drive reduction language, specifically, “an interference with the occurrence of an instigated goal-response at its proper time in the behavior sequence” [p. 7]). This frustration “built up” over different instances of frustration and continued to increase the probability of aggression. When aggression did occur, the frustration was simultaneously reduced (lowering the probability of immediate aggression) and the aggression reinforced (increasing the future probability of aggression). This reinforcement is due the fact that high frustration levels are aversive, and the lowering of such levels is rewarding.

Criticism of the Frustration-Aggression model came almost immediately on a number of different fronts. It was shown that frustration was not necessary for aggression. In fact Arnold Buss made a strong case that insult and personal attacks were much more likely to instigate aggression than was any kind of frustration (1961). Furthermore, “justified” frustration was shown to not be a cause of anger (Pastore,

1952; Cohn, 1955). Empirical problems with the posited cathartic effect of aggression were also reported (for reviews see Bramel, 1969; Geen & Quanty, 1977; Weiss, 1969). Despite these problems, the Frustration-Aggression model has persisted in the aggression literature, most likely because, as Tedeschi and Felson have said, “The theory was sufficiently precise as to allow experimental disconfirmations...” (1994, pg. 49). It is primarily because of its ability to generate testable predictions that it is included in this work.

#### Account of Features of Anger

Given that the magnitude of an imposed cost causes more frustration (a reasonable inference) the Frustration-Aggression Model clearly predicts that larger costs are more likely to cause anger and aggression. It cannot account for why the intention to cause harm is a better predictor of reactive aggression than is the harm itself however (see Cognitive-Neoassociationism above), and thus is inconsistent with the literature on intentionality. It is also inconsistent with the data showing that apologies drastically reduce aggression and anger, unless one posits that apologies somehow reduce frustration, though this does not follow from the original definition of frustration, namely “goal-blocking.”

The literature demonstrating that male-male aggression is typically fought over “face” or “status” is difficult for the Frustration-Aggression Model to account for. According to the Frustration-Aggression model, aggression is a means of removing a negative stimulus. To categorize another’s thoughts about you, or statements about

you in the case of insults, as a cause of negative affect stretches the concept (retrospectively) to fit any possible dataset.

Evidence of universality of some features is roughly consistent with the Frustration-Aggression Model, which argued that other species' aggression worked in similar ways; though it cannot account for the existence or design of universal facial expressions, vocal changes or physiological responses.

#### 5. A behaviorist account of anger (Arnold Buss)

##### Overview of Theory

Inspired by Thorndike and Skinner, Arnold Buss applied instrumental learning theory to the topic of human aggression (1961). His argument is that aggression is a learned behavior that is shaped and created through two primary types of rewards: pain cues on another individual (called *angry aggression*) and tangible rewards (*instrumental aggression*). He argues that instrumental aggression is the more common form of aggression and that it can be expressed in different ways (e.g. physical, verbal, active, passive, etc.) depending on the reinforcement history of each type of aggression.

Individual differences in one's tendency to aggress were theorized to stem primarily from a history of reinforcement for different kinds of aggressive behavior.<sup>8</sup>

Like most theories rooted in behaviorism, Buss's account of aggression is limited by philosophical problems that plague general purpose learning mechanisms of all

---

<sup>8</sup> Buss also posited effects for social facilitation of aggression and differences stemming from temperament and personality differences, but I will concentrate on the history of reinforcement as it is the primary cause of individual differences that stems from the logic of his theory.

kinds (see Cosmides & Tooby, 1987; Gallistel, 1990; Tooby & Cosmides, 1992).

While a review of these arguments is beyond the scope of this paper, the philosophical problems are sufficiently strong to rule out the possibility that positive reinforcement is sufficient to explain all aspects of aggression, at least in the absence of domain-specific anger mechanisms that restrict conditioning to certain stimulus situations and particular response options. The posited casual relationship between positive reward and increases in the frequency of aggression still remains as a potential explanation for some aspects of aggression though, such as individual differences in frequency of different types of aggression. It is logically possible that individual differences in aggression stem from differential reinforcement patterns – and this possibility will be further tested in Chapter 4.

#### Account of Features of Anger

Large costs should provoke reactive aggression because, to the extent that anger and aggression prevented or mitigated those costs, there would be negative reinforcement that increased the likelihood of future aggression. Behaviorism thus accounts well for the fact that larger costs more reliably provoke anger. It has the same difficulty as Cognitive-Neoassociationism (see above) when attempting to account for the role of intentionality, however. Behaviorism cannot explain the existence of a mental category such as “intentionality,” nor why such a category has an impact on whether reactive aggression is triggered. The effect of apologies and the importance of “face” and “status” are equally hard for Behaviorism to explain. Buss conceptualized verbal aggression as a form of “attack,” but unlike physical

attacks that lead to physical pain (a well-established punishment in behaviorist literature), it is unclear why verbal statements with little basis in reality would be considered aversive stimuli.<sup>9</sup>

The evidence of universality across cultures and cross-species neurological locality is difficult for learning theories to account for as it suggests that reinforcement patterns are extremely similar across cultures, individuals, and species. Furthermore, Behaviorism cannot account for the existence of universal facial displays of anger, changes in vocal patterns or specific physiological changes.

To be fair, it is not clear that Buss intended the theory to account for physiological changes resulting from anger. But it does point out the oddity of creating theories to account for the behavioral responses to anger without any reference to the process (i.e. natural selection) that designed the facial expressions, physiological causes, and cross-species universality in brain mechanisms that underlie human anger. It takes a very odd leap of faith to assume that natural selection carefully designed all these features of anger and then abandoned the most important aspect – behavior - to a simple learning mechanism that may or may not produce functional outcomes, and would be under the control of others who would not always have one's best interests in mind and would be capable of manipulating the mechanism by differential reinforcement.

## 6. Social Learning theory (Bandura)

### Overview of Theory

---

<sup>9</sup> It is an interesting feature of insults that they do not have to be true, or even plausibly true, for them

Bandura (1973, 1983) argued that Social Learning theory can account for many aspects of aggression and anger. Though not traditionally associated with rational choice theory, social learning theory posits many of the same processes (Bandura, 1977), including weighing the costs and benefits of behavior (often learned by watching the consequences visited on others as a result of their behavior). Bandura showed that by manipulating the consequences on others, children will mimic or not mimic the behavior exhibited by that other, including aggressive behavior. Much like a rational choice calculator, social learning mechanisms are theorized to track these reinforcers and respond based on the positive or negative valence of the likely outcome.

Bandura thus proposed that aggression is not the natural response to anger or frustration, but rather a behavior pattern that was reinforced because of the positive outcomes that typically result from aggression (particularly in childhood). Consistent with this, Patterson, Littman & Bricker (1967) showed that children who used aggression to terminate aversive treatment by other children increased their tendency to aggress. Their study also showed that aggression in school children is frequently reinforced (i.e. often led to positive outcomes), and that targets of aggression were seemingly chosen based on their tendency to not retaliate.

Because aggression is such a difficult variable to ethically elicit in the laboratory, many of Bandura's experiments (like most aggression experiments) had to use proxies of aggression, such as aggressive acts toward a life-sized toy doll (i.e. Bobo

---

to evoke anger.

doll). Such studies are limited by the fact that there is no real aggression taking place. The Patterson et al. study was naturalistic, but laboratory studies of aggression are necessary for demonstrating the effects of imitation as it is difficult to record naturally occurring imitation. But most laboratory experiments have used either pretend targets (e.g. Bobo dolls) or pretend aggression (e.g. shooting with a toy gun, hitting with a plastic mallet, etc.). See Tedeschi & Felson (1994) for a review of social learning experiments and their difficulty in demonstrating imitative aggression.

Conceptually, all reinforcement theories have difficulty explaining why some outcomes are reinforcing and others are not. One can either attempt to tie rewards and punishment back to known circuits of pleasure or pain, or surrender the explanatory burden to theorists interested in motivation. If one chooses the first option, then one has the difficult task of attempting to tie verbal insults into physical pain, or positive comments about one's choice of clothing into physical pleasure. If one chooses the second option, then the causes of anger and aggression will not be known until other theorists account for the nature of reward and punishment. Nonetheless, it is possible that individual differences in aggressive behavior are influenced by the rewards and punishments visited upon models that are then imitated.

#### Account of Features of Anger

Social Learning Theory, given its conceptual similarities to Behaviorism (see above), accounts well for the effect of costs but not the effect of intentionality on anger.

Like other “affect” models of aggression, Social Learning theory has a hard time accounting for why verbal indications of anything (i.e. apologies) reduce anger.

The role of “face” and “status” is more ambiguous. Violent aggression in non-police societies frequently led to injury or even death, which would have been readily observable to younger males. This presents a problem for Social Learning theory: the aggression in those cases was sometimes reinforced with increased status, but often led to torturous deaths. On the other hand, those who are physically aggressive often get status in these societies, and thus Social Learning theory could be made consistent with this data set – though it would have no explanation for what “face” is and why successful aggressive competitions lead to increases in it.

Again, the fact that individuals respond to insults with anger at all is perplexing from a number of different perspectives that require negative affect or vicarious exposure to negative affect to be the cause of aggression, such as Social Learning theory.

Social Learning theory, like Behaviorism (see above), is inconsistent with the cross-cultural and cross-species data on the universal features of anger.

## 7. Social Constructivism (James Averill)

### Overview of Theory

According to James Averill (1982, 1983), anger is best understood from a social-constructivist position which requires the endorsement of four major assumptions:

“First, emotions are responses of the whole person, and hence cannot be defined in terms of subclasses of responses (e.g., physiological or expressive reactions,

cognitive appraisals, instrumental acts, or subjective experience). Second, emotions are complex, polythetic syndromes; that is, no subset of elements or kind of response is a necessary or sufficient condition for the whole. Third, the rules that govern the organization of the various elements of the syndrome are primarily social in origin. Fourth, emotions serve a function within the social system, or at least are correlated with other behaviors that have a social function” (Averill 1983, pg. 1146).

With these assumptions as fundamental principles from which to theorize, Averill then goes on to discuss a series of self-report studies designed to answer six questions about the basic nature of anger (e.g. who do people typically become angry at, what are the consequences of anger, which gender experiences more anger). For each question he contrasts what typical anger theorists would predict (labeled *straw person position* because it is a conglomerate of theoretical positions) with what the person in the street would predict. There were no predictions derived or tested from the social constructivist position.<sup>10</sup>

The lack of predictive power inherent in social constructivism is sufficiently damning to prevent it from being taken seriously as a model of anger; but given the popularity of such approaches I will attempt as best I can to derive predictions that can be tested from a social-functional position.

The first and second of Averill’s assumptions, namely that emotions are responses of the whole person and no subset of elements are necessary or sufficient for the whole, does not lend itself to predictions other than that attempts to atomize anger

---

<sup>10</sup> This is true of Averill’s book (1982) as well.

will likely fail. This is difficult to understand as Averill, at the end of his 1983 article says, “One of the tasks of psychology should be to make such intuitive knowledge explicit, just as one of the tasks of linguistics is to explicate the rules (grammar) of the various human languages.” Such explication will be difficult if anger cannot be parsed into parts.

The third assumption is that the various elements are primarily social in origin. It is not clear what “social” means in this context, or which elements he is referring to. If one takes facial expressions of anger as one of the elements, then this assumption is already falsified, as facial expressions are universally the same for anger (Ekman & Friesen 1971). Without a theory that describes what effect society creates, there is no way to falsify or test this assumption.

The final assumption of the social constructivist position is that anger serves a function within the social system. The goals toward which social constructions should be designed are not clear - do they preserve the life of the individuals, their happiness, the longevity of all their ideas or just the ideas that distinguish them from others? Regardless, there can be some minimal and non-controversial predictions derived from such a position. It will be these predictions that are tested, when possible, in the studies described in Chapters 3, 4 and 5.

#### Account of Features of Anger

None of the known features of anger are explicitly or implicitly predicted by the Social Constructivist position. On the other hand, all of the evidence of universality across cultures and cross-species neurological locality contradicts a social-

constructivist position of anger. It is possible for a social constructivist to argue that the universal features of anger (e.g. facial expressions, physiological reactions, etc.) are part of a more “biological” system that has phylogenetic roots in common with other animals, but that the behavioral aspects of anger were left malleable. There are good reasons to believe that natural selection would not design organisms in this way because they would be ripe for manipulation by others who, almost certainly, do not have the individual’s best interests in mind.

#### 8. The Social Interactionist Perspective (Tedeschi and Felson)

##### Overview of Theory

James Tedeschi and Richard Felson advanced a comprehensive reinterpretation of the anger and aggression literature in 1994, calling for a common set of principles to account not just for aggression but of coercive behavior more generally. Their Social Interactionist Theory is explicitly functional, allowing one to derive predictions about the causes, computations and behavioral responses of anger – each designed (by rational choice) to bring about one of three possible functions: “(a) to influence others to obtain some benefit, (b) to express grievances and establish justice, and (c) to assert or defend social identities” (pg. 156). To accomplish these three goals, humans are theorized to “make ‘rational choices’ based on their perceptions of the value of outcomes, the probability of success in achieving those outcomes, the negative value of the costs, and the probability of costs for engaging in the contemplated act” (pg. 156).

Tedeschi and Felson's theory has difficulty accounting for some of the particulars of the goals that aggression and coercion are said to be used for – specifically a definition of “justified,” “norm,” and “social identities.” They rightly refer to literature demonstrating that male violence is frequently done to save “face,” but have no theoretical account for why males should be concerned with “face” or why physically beating another man restores this neurally-stored variable. They also use an intuitive definition of “unjustified,” making it difficult to derive predictions regarding what is considered justified and not. Though these parts of their computational model are left unspecified, the rational choice approach allows them to account for a great range of data while unifying the large dataset of aggression research under a few functions. Finally, this approach does not suffer from the “laundry list” problem as all three functions that are posited for coercive action are unified under a single principle of rational attempts to gain access to benefits.

#### Account of Features of Anger

The Social Interactionist Model predicts that imposed costs should cause anger to the extent that they indicate a lack of respect. This predicts and explains both the effect of large costs (assuming that imposition of a large cost does in fact show less respect than imposition of a small cost) and the effect of intentionality.

Though Tedeschi and Felson review the literature on apologies, it is not clear, from a rational choice perspective, why apologies exist, and why it is rationale for them to effect anger. They do not appear to restore a lost benefit, or serve the interests of justice (at least those apologies that do not involve restitution), so they would have to

be seen as ways of restoring the angry person's "social identity." This is possible, but hard to pursue computationally without a more specific theory of social identity.

The Social Interactionist Model explicitly refers to demeaning or insulting acts as a major cause of anger. It is not clear why physically beating another man should be the response to a demeaning or insulting act, or why these types of aggression are not practiced among women<sup>11</sup>. Finally, there is no computational account of what "face" "status" or "demeaning" refer to.

Insults are, according to Social Interactionism, predicted to be a primary cause of anger if they are attacks on one's social identity, but again there is no account of what characteristics should be part of one's social identity such that anger is triggered when they are contested (e.g. should it be insulting if someone suggests that one is short or tall, heavy or light, etc.)

Social Interactionism makes no predictions about the universality of anger.

## 9. The Regulator Theory of Anger

### Account for Features of Anger

Feature #1 - The Regulator Theory of Human Anger predicts that the cause of anger is not negative affect directly, but the indication that another holds a low Welfare Tradeoff Ratio with respect to you. A common indication of such a WTR is the imposition of a cost that is too large given the benefit received. Holding the

---

<sup>11</sup> It should be noted that Richard Felson predicted in 1982 that males should be more likely than females to attack an antagonist because "males may be more concerned with their identities in aggressive situations than females because of the traditional masculine identity stresses toughness and courage when one is attacked" (pg. 246). This explanation is circular (i.e. males are more violent because males have been more violent), not derived from a rational choice theory, and unable to

magnitude of the benefit received constant, the larger the cost one is willing to impose, the more likely anger is to be triggered.

Feature #2 – The Regulator Theory of Anger predicts that the perception that another holds a low Welfare Tradeoff Ratio should be a reliable cause of anger, and that this is indicated most clearly when someone imposes a cost of known magnitude for a benefit of known magnitude on an individual whose identity is known. While a fully described model of intentionality is beyond the scope of this paper, a starting point from which one could design such a model should include knowledge of: the magnitude of the cost they imposed, the magnitude of the benefit they received, and the identity of the individual on whom they imposed the cost. The third factor is described in more detail and tested in Chapter 3. One should also consider the possibility that intentionality is not a dichotomous variable such that all acts can be classified as “intentional” or “not intentional.”

Feature #3 - The Regulator Theory of Anger can account for the effect of apologies if they are viewed as explicit acknowledgements of a past discrepant Welfare Tradeoff Ratio that has been “fixed.” Such “fixing” may be best proved by restitution or by indicating a willingness to incur a cost to repay the angry individual.<sup>12</sup> The content of apologies are predicted to contain statements that

---

account for why this pattern is shown throughout all known human cultures (Brown 1991; Daly & Wilson 1988) and, in fact, all polygynous primates (Smuts et al. 1987).

<sup>12</sup> This theory of apologies differs from traditional theories in that it holds that a sincere apology will be one that indicates a more favorable WTR rather than simply restoring the cost that had been imposed on an individual. For example, an apology that motivates someone to impose a large cost on themselves to give you a minor benefit (indicating a high WTR toward you) will be seen as more sincere than an apology that motivates someone to impose a small cost on themselves to give you a large benefit. This contradicts a “restitution” view of apologies.

translate into the following cognitive grammar, “I will, from this point forward, demonstrate a more favorable Welfare Tradeoff Ratio with respect to you, such that I will no longer impose costs of that magnitude on you for benefits of that magnitude.”

Feature #4 - The Regulator Theory of Human Anger suggests that, like other animals, misperceptions of formidability play a part in the setting of Welfare Tradeoff Ratios and positions in dominance hierarchies. As such, demonstrations of physical strength in humans should involve the same procedures as strength contests in non-human animals, such as: signals of challenge, escalating conflicts that start with low-violent demonstrations (e.g. pushing contests, staring contests, etc.) and escalate to more violent demonstrations (e.g. wrestling, punching, weaponry), and signals of surrender that cease violence. These patterns appear to fit the data on homicide quite well (Luckenbill, 1977). Only the Regulator Theory of Anger predicts this style of aggressive interaction explicitly. Furthermore, it is predicted that these contests should be more common among males than females because humans are mildly polygynous, and in polygynous species males are the more aggressive of the species (Daly & Wilson, 1983).<sup>13</sup>

Feature #5 - Just as other animals have clear signals of aggressive intent, so are humans predicted to have low-cost means of i) provoking a conflict of assessment (e.g. a fight over who is “tougher”) and ii) indicating disagreement about a resource division. At least some data gathered on the content of insults is consistent with this

---

<sup>13</sup> Sex differences in aggression are one of many of cluster of theoretically-related variables that track the magnitude of polygyny across species and include: bimaturationism (males develop slower), sex differences in death rates (males die soon), sexual dimorphism (males are larger) and testicle to body

Regulator Theory of Human Anger – which predicts that anger-inducing insults can be understood as attempts to directly indicate a low Welfare Tradeoff Ratio with respect to another individual, and, perhaps, as attempts to influence others to hold lower WTRs with respect to the insulted individuals. Insults should, psychologically, translate to the form, “I do not value your interests highly.” More proximately, insults can be declarations of a deficit in a variable that is used to determine one’s Welfare Tradeoff Ratio toward that individual. One such variable, for males at least, is physical strength. Fitting with this theory, insults toward males have been shown to frequently target a man’s strength - e.g. wimp, wuss, geek, nerd, bitch, girly-man, pussy, weakling (Harris 1993; Preston & Kimerley 1987). Such insults are almost never leveled at women (when “bitch” is used toward women it indicates something about character not physical strength). Though beyond the scope of this analysis, it seems likely that other insults fit into categories that make a man socially powerful or not, such as his intent to cooperate (e.g. asshole, prick, bastard, rat etc.), his being unable/unwilling to function as a reliable cooperator or being otherwise unworthy of having others take his interests into account (e.g. punk, white trash, ghetto trash, bum, etc.), or his competence (e.g. idiot, fool).

Feature #6- In contrast to other theories, the Regulator Theory of Anger clearly predicts that anger is an adaptation designed by natural selection and that its basic computational structure should be universal across cultures and should share a phylogenetic relationship with closely related non-human animals resulting in

---

size ratio (polygynous species have larger testicles). On all these variables humans are where they are

similarly localized brain areas. Furthermore, like non-human animals, humans should have signals of resource contention (e.g. facial and vocal expressions), physiological preparedness for aggression if necessary, and a structured functional set of causes and behavioral responses that are similar across cultures.

*C. Summary of theories and their ability to account for features of anger*

Table 2-1 below summarizes the ability of various models of anger to account for the robust and powerful common features of anger discussed above. The following coding scheme was used: (2) – the theory in question clearly predicts the feature and the absence of the feature would be evidence against the theory; (1) – the theory is consistent with the data point but does not require it; (0) – the theory makes no prediction about feature; (-1) – the theory is contradicted by the evidence for that feature of anger.

---

predicted to be based on the degree of polygyny.

**Table 2-1: Predictions of current models against current data<sup>14</sup>**

	Costs lead to anger	Intent. More important than cost	Apologies	“Face” – saving aggression	Insult	Universal features across cultures <sup>15</sup>
Regulator Theory	2	2	2	2	2	2
Cognitive Motivational Relational	2	2	1	2	2	1
Cognitive Neoassociation	2	-1	-1	-1	1	1
Equity Theory	2	1	1	1	-1	1
Frustration Aggression	2	-1	-1	-1	-1	1
Behaviorism	2	-1	-1	-1	-1	-1
Social Learning	2	-1	-1	1	-1	0
Social Constructivism	0	0	0	0	0	-1
Social Interactionism	2	2	1	2	1	0

<sup>14</sup> I have tried to be as fair as possible when coding the ability of these theories to predict these features (see above for explanations); but I should reiterate that different features would no doubt be picked by the individuals who penned these theories.

<sup>15</sup> This column should be viewed as pertaining to the universality of antecedents and physiological responses. None of the theories can account for the existence of facial expressions, vocal changes, or other features clearly designed by natural selection in ways unconnected to the mechanisms posited in the various theories.

## Chapter 3

### Study Set A – the Causes of Anger Elaborated

“At present we have no unified theory of the causes of human anger” –Russel G. Geen (1998) pg. 317  
“Arguments from evolutionary biology have generally been ignored or dismissed by social psychologists” –Russel G. Geen (1998) pg. 319

This chapter describes four experiments that were designed to test predictions derived from the Regulator Theory of Anger regarding the causes of anger.

Experiment A1 tests the prediction that, holding cost imposed constant, less anger is elicited when the offender receives a high rather than low benefit – in other words when the value of a resource to the offender ( $V(Y)$ ) is high.

Experiment A2 tests the prediction that anger is increased when the offender has knowledge of the angry person’s identity at the time they impose the cost, and the angry person knows this.

Studies A3 and A4 further test the effect of the offender’s benefit on anger as well as explore how individuals can avoid triggering anger in others by using arguments designed to contradict the notion that they have a low Welfare Tradeoff Ratio.

***A. Experiment A1: Is less anger elicited when a cost is imposed to obtain a high benefit than when the same cost is imposed to obtain a low benefit?***

The most direct indicator of a low Welfare Tradeoff Ratio is when someone knowingly imposes a cost for a relatively small benefit; for example, someone taking something that is much more valuable to you than to them, or withholding something that is of particular value to you even though it is not particularly valuable to them.

Two simple hypotheses derived from the theory are as follows:

1) All else equal, the larger a cost imposed on an individual the more likely that individual is to become angry.

2) All else equal, the smaller the benefit reaped as a result of imposing a given cost on an individual, the more likely it is that the individual will become angry.

Prior evidence, reviewed in Chapter 2, confirms the prediction that, all else being equal, the larger the cost imposed on an individual, the more likely it is that anger will be triggered.

Much less theoretical work has been devoted to determining the effect of the offender's valuation of the resource (i.e. the benefit reaped by the individual who imposed the cost). Different theoretical perspectives on anger make different predictions about the role of the benefit a person receives as a result of imposing the cost; they are summarized in Table 3-1. It is because of these differing predictions that this effect was tested.

**Table 3-1**

**Experiment A1 - Predictions from multiple perspectives**

<b>Theoretical Perspective</b>	<b>Effect of Increasing Offender's Benefit on Anger</b>
Regulator Theory	Decreases anger
Cognitive-Motivational-Relational	Unspecified
Cognitive Neoassociationism	Unspecified
Equity Theory	Increases anger
Frustration-Aggression	Unspecified
Behaviorism	Unspecified
Social Learning theory	Unspecified

Social Constructivism	Unspecified
Social Interactionism	Unspecified
Folk Deterrence	Increases anger

The Regulator Theory of anger predicts that the inference that someone’s WTR toward oneself is low is a specific trigger for anger. When an offender is willing to impose a given cost on you for a small benefit, that indicates that the offender is not taking your welfare much into account—that the offender’s WTR toward you is low. In contrast, willingness to impose the same cost for a very large benefit does not allow you to infer that the offender’s WTR toward you is low, because s/he would impose that cost-benefit transaction even if his/her WTR toward you is high. This analysis predicts that, when all else (including cost imposed) is held equal, an offense that results in a small benefit to the offender will trigger more anger than one that results in a large benefit to the offender.

Equity theory makes a contrasting prediction. It suggests the function of anger is to realign the outcomes of an interaction so as to arrive at an equitable distribution. Those who are “under benefited” inflict costs on those who are “over benefited” in order to even out the distribution of the resource. This predicts that the larger the benefit a person gains as a result of imposing a cost, the more angry the victim should be, as such a distribution would require not only enough anger and punishment to even out the imposed cost, but also enough to even out the benefit (e.g. if someone imposed a cost of 5 units in order to receive 5 units, it would take 10 units of punishment to even the score; while if they imposed the same cost to receive 1 unit, it

would take only 6 units of punishment to even the score.) The Equity Theory approach to human anger would thus predict a positive relationship between anger expressed by the victim of an offense and the benefit the offender received as a result of the offense: all else equal, the more someone benefits by doing you wrong, the more inequitable the outcome and, therefore, the angrier you should be.

The other theories reviewed are silent on how the magnitude of the offender's benefit should regulate anger. Cognitive-Motivational-Relational theory describes the core cause of anger as a demeaning offense against an individual (Lazarus, 1991), but leaves unspecified the computational nature of "demeaning." Lazarus suggests that an act cannot be demeaning unless the act is blameworthy, defined in part as a person's ability to control their own actions. "If...the people who frustrated us are capable of having acted otherwise, then they are likely to be blamed," (pg. 223). But again there is no computational specification of the variables that individuals' use to determine whether or not someone could have acted differently. Without such a model we cannot derive specific predictions about the role of the offender's benefit on anger from cognitive-relational theory.

The Cognitive Neoassociationism view of human aggression is that anger results from a complex interplay between negative affect and a variety of cognitions that interpret that affect (Berkowitz & Heimer, 1989; Berkowitz, 1990). It makes no specific predictions about the role of the offender's benefit.

The traditional Frustration-Aggression model (Dollard et al., 1939) predicts that aggression is determined by the magnitude of frustration, and does not reference the

motive of the instigator of that frustration as a useful variable. Thus it predicts no effect of the magnitude of the benefit to the offender on the anger of the victim.<sup>16</sup>

Behaviorism, Social Learning theory, Social Interactionism and Social Constructivism make no explicit predictions about how the magnitude of the offender's benefit regulates levels of anger.

Though not often made explicit, there is a folk notion that anger functions to motivate aggression which deters an individual from doing something that was costly to the angry person. Models of justice and the social function of punishment also suggest that socially-sanctioned aggression sometimes serves to deter crime (Fridja, 1994; Jacoby, 1983). One could argue, in line with the folk notion of deterrence, that anger serves as a motivation to enact punishment that would make it no longer in the best interests of the offender to commit the crime or anger-provoking act.<sup>17</sup> A simple prediction from such a theory would be that the magnitude of punishment (and thus the amount of anger necessary) to deter the offender will increase with the benefit to the offender. If someone takes \$500 then one must deliver \$501 worth of punishment while if someone takes \$50 then one requires only \$51 worth of punishment. Thus a

---

<sup>16</sup> It is worth noting that early publications on the FA model made mention of an instigating factor dubbed "arbitrariness." Offenses that were seen as arbitrary were said to increase the frequency of anger (Cohn, 1955). Arbitrariness was operationalized in ways that conflated the offender's benefit, intentionality, and unexpected behavior (e.g. a bus driver not doing his job and stopping at a bus stop). It should also be noted that the role of "arbitrariness" does not flow from the logic of the Frustration-Aggression Model and seems to have been added on to the theory when evidence suggested frustration was not the only instigating factor in aggression. It is not clear that the theory should be credited with making a prediction about arbitrariness.

<sup>17</sup> Fessler and Haley (2003) for example, argue that, "anger motivates actors to strike out at those who transgress against them, thus inflicting costs on the transgressor which then reduce the attractiveness of future attempts at transgression. The stronger the response to transgression, the greater the deterrent effect..." (pg. 13)

folk deterrence model of anger, like Equity Theory, would predict that anger increases as an offender's benefit increases.

I used a simple between-subjects experiment with a vignette to test the hypothesis that subjects will be more angry when a contested resource is of low value to the offender than when the contested resource is of high value to the offender. In both conditions a classmate cuts in line in front of the subject at a pay phone in order to redeem a lottery ticket on time. The value of the ticket was varied between conditions. In the high benefit condition it was worth \$1000. In the low benefit condition it was worth \$5. Everything else, including the cost imposed, was held constant across conditions. Moreover, the size of the benefits were chosen relative to the cost imposed; that is, relative to the cost imposed (having to wait an hour), local informants viewed \$1000 as a large benefit and \$5 as a small one. Note that if the cost imposed were much higher—breaking a leg rather than waiting an extra hour--\$1000 would not necessarily count as a “large” benefit. In a case such as this, one could not unambiguously predict a lessening of anger in the \$1000 because when the cost is as high or higher than the benefit, the cost-benefit transaction may still represent a low WTR toward the subject.

### **Method for Experiment A1**

Participants: 109 undergraduates from UCSB participated in this experiment for partial fulfillment of their introductory psychology class credit. There were 37 males (mean age 18.4, ranging from 17 to 20) and seventy-two females (mean age 18.2,





**Prediction 1.** After the offense was committed, but before learning about the lottery ticket, subjects should feel angry (manipulation check).

**Prediction 2.** After learning the value of the lottery ticket, subjects should be less angry in the high benefit condition. In this condition, the benefit gained is large relative to the cost imposed, so the action reveals little about the offender's WTR.<sup>18</sup>

**Prediction 3.** The diminution in anger should be greater in the high benefit condition than in the low benefit condition.

Prediction 1 is not unique to the regulator theory; Predictions 2 and 3, which involve changes in anger, are. Two dependent measures were used to assess changes in how angry the subject felt, as described below.

### **Results and Discussion for Experiment A1**

*Manipulation check and preliminary analyses - Did subjects report that they would feel angry after being pushed out of the way, but before they learn why the classmate did this?*

Yes. The mean anger rating was 6.44 (SD .698) out of 7 in the high benefit condition, and 6.14 (SD .833) in the low benefit condition. These ratings were made before the independent variable (the benefit manipulation) was introduced. These ratings confirm prediction 1, and demonstrate that the situation described in the text was sufficient to elicit anger.

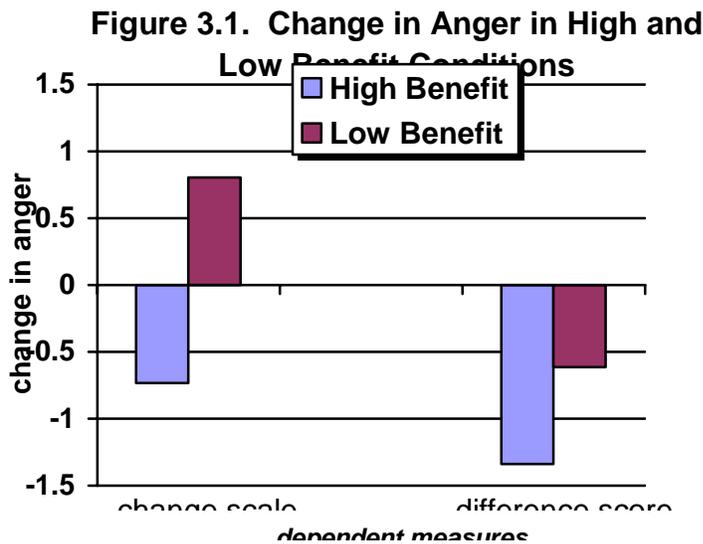
---

<sup>18</sup> Note that there is no prediction about whether the LB condition will cause an absolute increase or decrease in anger. Anger should decrease if subject's prior assumption was that the offender gained nothing by pushing them aside; it should increase if they think \$5 is a trivial sum and had assumed the offender would have had to gain more to justify pushing them aside.

Were levels of anger before the benefit manipulation the same across the high and low benefit conditions?

No. An independent samples t-test revealed that those in the HB condition rated themselves higher on the pre-benefit anger measurement than those in the LB condition ( $t(107) = 2.041, p = .044, r = .19$ ). As this measure was taken *before any* manipulation, this difference must reflect random variation in subjects' propensity to anger.

For this reason, it is particularly important to use the subject as his or her own control. The change in anger scale does this. Immediately after learning about the lottery ticket, subjects were asked, *Compared to your earlier answer, how much anger would you feel now?* with a scale ranging from much less anger (-3) to much more anger (+3).



Was anger reduced in the high benefit condition? Was it reduced more than in the low benefit condition?

The answer to both questions is yes (see Figure 3.1). The change scale results show that subjects in the high benefit condition reported a lessening of anger ( $M = -.731$ ,  $SD = 1.36$ ), whereas those in the low benefit condition reported a heightening of anger ( $M = .807$ ,  $SD = 1.32$ ). To determine whether anger was reduced more in the high benefit condition than in the low benefit condition, an independent samples t-test was conducted. The difference between these conditions was significant ( $t(107) = -6.00$ ,  $p < 0.001$ ), with a moderately high effect size ( $r = .50$ ). This effect size represents a substantial diminution in anger in the high benefit condition relative to the low benefit condition.

After filling out the change scale, subjects had rated their emotional experience on six different emotions, including anger. Although it is a more indirect measure, a different way of assessing change in anger is to create a difference score for each subject by subtracting their pre-benefit manipulation anger rating from their post-benefit manipulation anger rating. (This is necessarily a noisier measure; one might experience a clear reduction in anger, but if the initial number picked is not remembered with precision, this later scale might not reflect that reduction.)

Like the change scale, the results for the difference score measures also address Predictions 2 and 3. By testing the difference score for the high benefit condition against the hypothesis that there was no change in anger (0), we find that Prediction 2 is confirmed: Subjects in the high benefit condition experienced a significant reduction in anger, with a large effect size of  $r = .69$  ( $M = -1.34$ ,  $SD = 1.4098$ ;  $t(51) = -$

6.836,  $p < .001$ ).<sup>19</sup> By this criterion, subjects in the low benefit condition also experienced a reduction in anger, with an effect size of  $r = .50$  ( $M = -0.6140$ ,  $SD = 1.0816$ ;  $t^*(56) = -4.286$ ,  $p < .001$ ). (This diminution in anger for the low benefit condition is different from the increase found using the change scale, but there is no way to tell whether this is due to noise in the measure or dissipation in anger as the subject continues to answer questions. Even if dissipation occurs, however, this would affect the high and low benefit conditions equally.)

Importantly, Prediction 3 is confirmed using the difference score: The reduction in anger was greater for subjects in the High Benefit condition than for those in the Low Benefit condition ( $t(107) = -3.02$ ,  $p < .005$ ,  $r = .28$ ). Thus both measures of change in anger—the change score and the difference score—supported Prediction 3: Learning the offender imposed a cost to obtain a large benefit reduced anger more than learning the cost was imposed to obtain a small benefit.

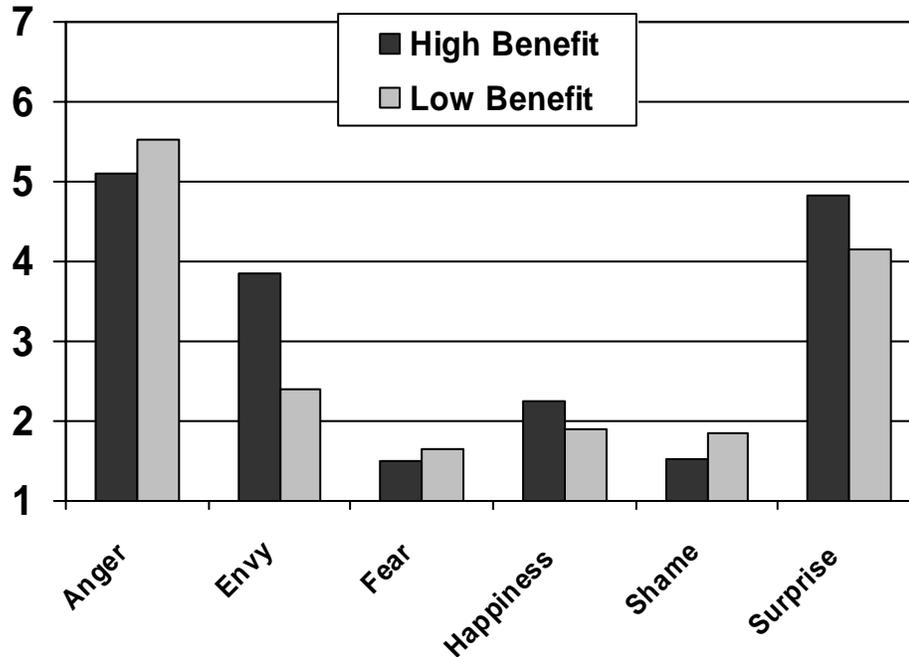
There were no significant sex differences predicted or found.

Figure 3-2 shows the means of other emotion questions for each condition.

---

<sup>19</sup> There is a debate about the proper way to calculate effect sizes in paired samples data. I follow Rosenthal (1991) and compute the effect size from the paired samples  $t$ , though some argue that this leads to an inflation of the effect size, see Dunlop et al. (1996).

**Figure 3-2: Self-report of Emotions when offender received High Benefit (HB) or Low Benefit (LB) as a result of imposing a cost.**



The change scale and difference scores used the subject as his or her own control, which was necessary given the preliminary analysis showing that the mean anger scores taken before the benefit manipulation were not the same in the HB and LB conditions: the high benefit subjects were a slightly angrier group. When one does not control for this difference by using within-subject measures, and merely compares the final self-report of anger of those in the high and low benefit groups, one still sees less anger in the high than the low benefit condition, but the effect is more muted ( $t(107) = -1.589, p=0.11, r = .15$ ). This reduction in effect size is to be expected given that the high benefit subjects had higher anger levels independent of the benefit manipulation.

As can be seen on Figure 3-2, those in the High Benefit condition differed on measures of envy and surprise as well as their reported changes in anger. This opens up a number of possible confounds that could explain the relationship between the value of the resource to the offender and the subject's change in anger.

*Is the change in anger the result of differential envy?*

It may be reasonable to assume that humans are less prone to be angry with those who are capable of providing valuable resources to them, to thus avoid the proverbial biting of the hand that feeds them. As such, it may be possible that subjects in this experiment had a lessening of anger because of an increase in envy and not because of an implicit calculation of the Welfare Tradeoff Ratio used by the offender. To test this hypothesis a partial correlation was computed between the benefit to the offender (i.e. \$1000 or \$5) and self-reported change in anger while controlling for self-reported envy. The effect of the manipulation on the change in anger remained, partial  $r = 0.45$ ,  $p < .001$  (simple  $r = 0.50$ ). It is concluded that the differential changes in anger in the High Benefit condition and Low Benefit condition were not mediated by envy.

*Is the change in anger the result of differential surprise?*

It may also be reasonable to assume that humans reduce their anger in situations that they do not fully understand, so as to avoid dangerous overreactions in situations that otherwise would not require anger. Thus surprise may have a suppressing effect on the intensity of anger, at least until more information is gathered. If this is true, anger may have lessened in the High Benefit condition, not because of the implicit calculation of a Welfare Tradeoff Ratio, but because there was an increase in surprise.

To test this hypothesis a partial correlation was computed between the benefit to the offender and change in anger while controlling for self-reported ratings of surprise. Results show that the effect of the offender's benefit remained, partial  $r = .49$ ,  $p < .001$  (simple  $r = .05$ ). It is concluded that the differential change in anger was not due to differences in surprise.

*How do we know subjects were making an implicit calculation of the Welfare Tradeoff Ratio used by the offender? Qualitative data.*

To gain insight into the decision process the subjects were using when reporting their change in anger, after rating all the emotions subjects were asked to report why they were initially angry and then why their anger level changed (if it did in fact change) after learning the value of the ticket. Though implicit computations are sometimes difficult to make explicit, those subjects whose self-reported change in anger decreased in the High Benefit condition gave reasons that hinted at or made explicit the calculation of a Welfare Tradeoff Ratio.

Typical examples from subjects whose anger was reduced in the High Benefit condition include: "His need for the phone was far greater, \$1000 versus me waiting an hour isn't a big deal;" "I was angry because she shoved me rudely out of the way and assumed her call was more important than mine. It changed because her situation ended up being slightly more important than mine;" "It changed because she had a more legit reason than I would to be pushing to the front of the line;" "why is her problem/need to use the phone stronger than mine. It changed because losing a

thousand bucks would not be fun and I guess it would stink if she called two minutes too late.”

Note the repeated reference to the relative valuation of the phone; in particular that the individual who took the phone from the subject valued it more highly.

It is interesting to compare these responses to those given by subjects in the low benefit condition who became more angry when they learned the ticket was for only \$5. These include: “He caused me a one hour wait for a \$5 lottery ticket, and from his dress he is probably rich;”<sup>20</sup> “it was frustrating to know I missed the bus over a \$5 lottery ticket”; “It increased a little because I have to wait another hour for the bus so that he can gain \$5”; “It changed and got even stronger when the subject was about a measily \$5.00 lotto ticket”; “It was only a \$5 ticket. It wasn’t something more urgent than me missing my ride;” “I was angry because his call is not any more important than mine”; “It changed because the reason she pushed me was not very important. It was not life or death and it wasn’t very much money;” “I was a little more mad afterwards because winning five dollars in the lottery is nothing, I could have given her that money. If her phone call was an emergency I wouldn’t have cared, but it wasn’t.”

---

<sup>20</sup> In the vignette the other character was referred to as a “well-dressed guy” or “well-dressed girl,” so that subjects wouldn’t assume in the Low Benefit condition that the character was poor and in the High Benefit condition that the character was wealthy. Had subjects made that assumption they could weight the Low Benefit character’s valuation of the \$5 ticket as comparable to the High Benefit character’s valuation of the \$1000 ticket. Thus there would have been no manipulation of the offender’s change in welfare.

Again note the references to the relative valuation, in particular that the value of the ticket was not greater to the individual than to the subject. Also notice that several of the subjects spontaneously reported that they would not have been as mad if the individual valued the phone more.

The written reports from those subjects whose change in anger went in other directions were also telling. The following examples are from subjects whose anger increased or didn't change despite learning of the \$1000 ticket. "She was only pushing me out of the way to secure a monetary reward. It is wrong for her to think that \$1000 is more important than my life/feelings." "I would be angry with her because she felt that her phone call was more important than my phone call. She decided that she was more important than me. I became angrier after reading the second part because ... there was no valid explanation for the girl's rudeness, thus I would become angrier at this person." Note the explicit reference to the trade-off ratio "\$1000 is (not) more important than my life/feelings," and "she felt her phone call was more important than my phone call." Also note in the second example that the subject concludes that the girls' acceptance of this cost-benefit transaction indicates that she believes herself to be more "important" than the subject. This was frequently noted by subjects, such as in the following example, "I was angry with her because it is always rude to cut in line and shove someone out of the way, as if you are more important...". In the high benefit condition, there were only 4 out of 52 subjects whose anger increased, and 14 whose anger did not change.

Only 6 of 57 subjects in the Low Benefit condition reduced their anger (change scale). One wrote, “I was less angry because I understood her urgency.” Another apparently inferred that the lottery ticket was of great subjective value even if it was only \$5, “You really don’t know if it was important to him. He could really need even a small amount of money.”

Thus based on the subjects’ written descriptions of their thought process, it seems clear that they were calculating a cost-benefit ratio and reserving anger for situations in which someone imposed a cost on them for what they perceived as a relatively low benefit.

**Summary and Conclusion for Experiment A1.** The regulator theory of anger proposes that the perception that an offender has a low WTR toward you is a specific trigger for anger. Imposing a cost for a high benefit does not reveal a low WTR, whereas imposing that cost for a low benefit does. Hence, discovering that the offender harmed you for a high benefit should reduce your anger compared to discovering s/he harmed you for a low benefit. The results of Experiment A1 support this prediction. Equity theory predicts more anger in the high benefit condition (where there is more inequity), and so does the folk theory discussed. Thus the results, showing less anger in the high benefit condition, are the opposite of what they should be if those theories were correct.

***Experiment A2: Is anger mitigated when the offender did not know the angry individual’s identity?***

According to the Regulator Model of human anger, anger is triggered by the perception that another holds a low Welfare Tradeoff Ratio with respect to the angry individual. Imposition of a cost without knowledge of the victim's identity, however, would not indicate the Welfare Tradeoff Ratio used, as it would have been impossible for the offender to have computed or accessed a stored WTR without knowing the victim was in fact the one on whom the cost was imposed. The regulator theory therefore predicts less anger when the offender did not know the victim's identity than when s/he did not.

None of the alternative theories of human anger reviewed in Chapter 2 have explicit theories of intentionality that would make predictions about the importance of the offender's knowledge of victim identity. It is possible, however, to see how Cognitive-Motivational-Relational theory, Social Interactionism and Equity Theory would predict that offenses would not be as "demeaning" "identity threatening" or indicative of a long-term inequitable relationship. For the other six alternative theories, there is no theoretical rationale for why the offender's knowledge of the victim's identity should regulate anger.

A small vignette about a practical joke was used to determine the effect of the offender's knowledge of victim's identity on anger elicited in the victim. The joke involved an insult, with a potential loss of face in front of others.

### **Methods for Experiment A2.**

Participants: Sixty-four undergraduates from UCSB participated in this experiment for partial fulfillment of their introductory psychology class credit (21 male (average

age 18.4, ranging from 17 to 20), 43 female (average age 18.4, ranging from 17 to 21). They were randomly assigned to one of two conditions, with 30 in the Known Victim condition and 34 in the Anonymous Victim condition.

Procedure. Subjects were brought into the lab in groups of two to four and were handed the questionnaire which contained a vignette about a person who puts a slug in the subject's lunch during a class biology trip. In the Known Victim (KV) condition, the person knew whose lunch it was; in the Anonymous Victim (AV) condition the person did not know it was the subject's lunch at the time of the act.

Materials specific to Experiment A2. The following is the Known Victim vignette. The character in the vignette was always male. As in Experiment A1, there was a page break prior to "The story continues here" so that subjects would rate their initial anger before reading the manipulated section of the vignette. Thus subjects could, as before, serve as their own controls.

You and a friend sign up for a biology class here at UCSB that requires you to go out with the class and observe wildlife in the forest. The Saturday trips take place between 10:00am and 3:00pm, so students bring bagged lunches along. The lunches are kept in the bus so they don't distract the wildlife.

You get to know the other students on the trip, but everyone pretty much stays with their lab partner. Everyone got to pick their lab partners, so you and your friend got to work together.

One student, Jimmy, is a bit of a prankster and occasionally tells an offensive joke or two. One day when everyone gets back to the bus, the TA opens up the bin the lunches are kept in, and you see one of the lunches has been marked with a black magic marker. It reads: "At last, meet your intellectual equal." The TA notices the bag moving, and when he opens it a four-inch banana slug crawls out of it. Everyone cracks up laughing. One of his friends congratulates Jimmy on the joke. "Good one," he says. "Thanks," Jimmy replies.

It is your lunch. Everyone watches while you get your lunch and throw the slug aside.

**Please answer the following question before continuing on with the story:**

How angry would you be feeling at that moment? (circle a number)

1      2      3      4      5      6      7

not at all

very strongly

**\*The story continues on the next page\***

[Page break]

**The story continues here:**

As people continue to laugh at you, one of the other students whispers to Jimmy, “Did you know whose lunch it was?” Jimmy responds, “Yes.”

In the Anonymous Victim condition the last word was changed from “Yes” to “No.”

All other aspects of the vignette were identical. Subjects then reported how much their anger changed on a Likert scale from -3 (much less) to +3 (much more). Finally they reported their emotional experience on Likert scales from 1 (not at all) to 7 (very strongly) for the following emotions: surprise, anger, sadness, fear, envy and shame.

### **Predictions of the Regulator Theory of Anger**

Regardless of their pre-manipulation levels of anger, subjects who discover that the insult was directed specifically at them should be more angry than those who discover it was not.

### **Results and Discussion of Experiment A2.**

*Manipulation check and preliminary analyses - Do subjects report that they would feel angry after the slug was pulled out, but before they learn whether the prankster was targeting them personally?*

Yes. The mean anger rating was 5.33 (SD = 1.39) out of 7 in the Known Victim condition, and 4.97 (SD = 1.64) in the Anonymous Victim condition. These ratings were made before the independent variable (the knowledge manipulation) was introduced. These ratings confirm prediction 1, and demonstrate that the situation

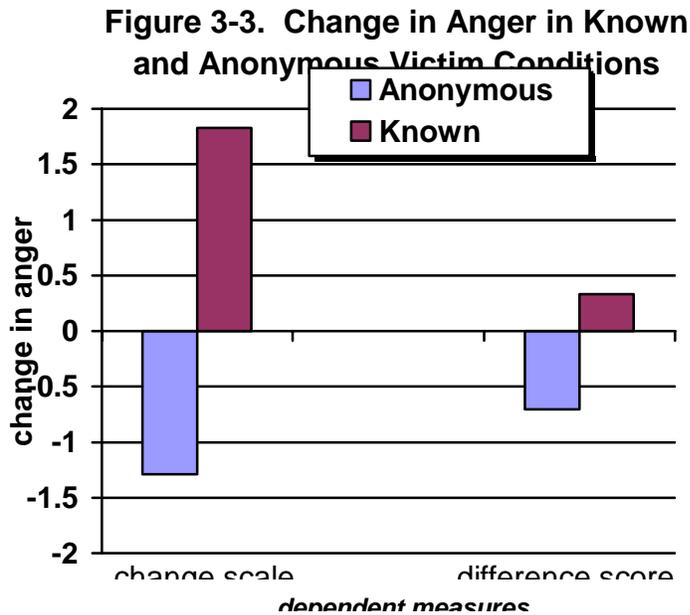
described in the text was sufficient to elicit anger.

*Were levels of anger before the knowledge manipulation the same across the known and anonymous victim conditions?*

Yes;  $t(62) = -.945, p = .348$ ). Nevertheless, using the subject as his or her own control remains a more sensitive test. Thus the analyses were done on the change scores and difference scores, as in Experiment A1.

*Is anger mitigated by discovering that the offender did not know your identity?*

Yes. This is true whether change in anger is assessed using the change scale or the difference score measure (see Experiment A1 for discussion). See Figure 3-3.



The change in anger scale, which runs from -3 (much less angry) to +3 (much more angry), revealed different results for the Known Victim and Anonymous Victim conditions. In the Known Victim condition, the subject discovers that s/he was the intentional target of the insult/joke. This discovery increased their anger ( $M = +1.83$ ,

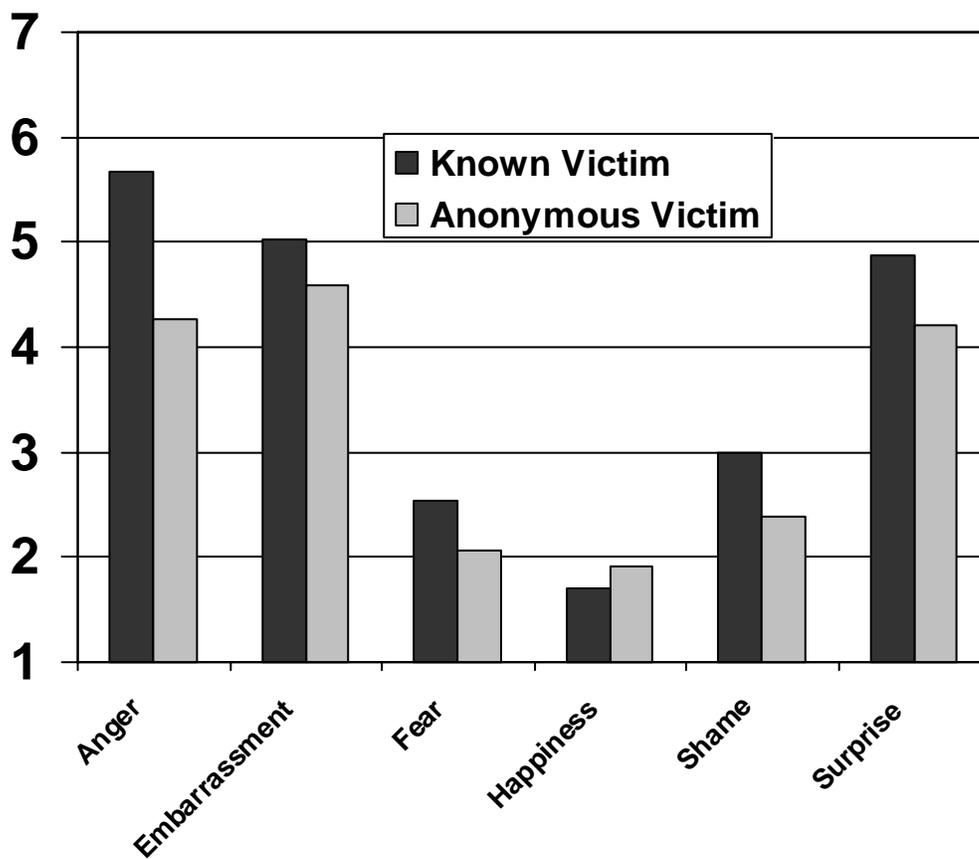
SD = 1.36). In contrast, anger decreased in the Anonymous Victim condition, after the subject discovered that the prankster did not know who the ultimate target of the insult/joke would end up being ( $M = -1.29$ ,  $SD = 1.14$ ). The difference between the Known and Anonymous Victim conditions was significant ( $t(62) = -9.97$ ,  $p < 0.0001$ ), and each condition was significantly different than zero, KV:  $t(29) = 7.35$ ,  $p < 0.0001$ ; AV:  $t(33) = -6.61$ ,  $p < .00001$ . As predicted, subjects were less angry when they learned the insult/joke was not directed at them personally than when they learned that it was. The size of this identity effect was large:  $r = .785$

The same pattern emerges when one uses the (noisier) difference score, in which each subject's pre-manipulation anger rating is subtracted from their final one. In the Known Victim condition, the mean difference score was significantly greater than zero (indicating more anger after learning the prankster knew who he was insulting), ( $M = .333$ ,  $SD = 0.80$ ),  $t(29) = 2.28$ ,  $p = .03$ . In contrast, when subjects in the Anonymous Victim condition learned the prankster had not been targeting them specifically, their anger decreased significantly ( $M = -0.706$ ,  $SD = 1.22$ ;  $t(33) = -3.376$ ,  $p = .002$ ). Most importantly, anger decreased more in the anonymous than in the Known Victim condition ( $t(62) = -3.969$ ,  $p < .001$ ,  $r = .45$ ), showing that subjects become less angry when they learn the insult was not intended specifically for them than when they learn that it was.

Significant sex differences in anger were neither predicted nor found. Females reported significantly more fear ( $p < .05$ ) and sadness ( $p < .05$ ) than males did. Neither variable interacted with condition, however, and thus these differences cannot explain

the effect of the manipulation on anger. Final emotion ratings are summarized in Figure 3-4 below. Final anger ratings differed between conditions ( $t(62) = -3.682, p = .0005$ ), but the other emotions yielded no significant difference between conditions.

**Figure 3-4: Self-report of Emotions when offender knew and didn't know the victim's identity**



*Was the change in anger due to differential embarrassment?*

The Regulatory Theory of anger, as well as other theories, suggests that high costs imposed on the victim will result in more anger than lesser costs. Thus it becomes crucial, when testing for causes of anger, that one hold the cost imposed on the victim constant. This was done in Experiment A1 by changing a single word in the story

(thousand versus five), one that could not reasonably be expected to change the cost imposed on the subject (having to wait an hour for the next bus). Similar efforts were taken in Experiment A2 to only manipulate the story in the slightest way so as to change the knowledge of the victim's identity and nothing else. Despite these efforts, it is possible that when the character in the vignette disclosed to his friend that he knew or did not know the owner of the lunch, subjects in the KV condition assumed that more people knew about their predicament, thus elevating their level of embarrassment (the cost being imposed).

This does not appear to be the case, however, as subjects in the KV condition did not differ significantly on questions about shame or embarrassment. There was, however, a non-significant relationship, with subjects in the KV condition being more embarrassed and feeling more shame. To ensure that this difference in embarrassment (as a proxy for cost) was not the cause of the differences in the change in anger, knowledge of victim's identity was regressed onto the change in anger, controlling for reported embarrassment and shame. Results indicate the effect of the offender's knowledge of the victim's identity was still strong, partial  $r = .781$ ,  $p < .001$  (simple  $r = .785$ ). It is concluded that increases in shame and embarrassment did not mediate the change in anger.

*Are subjects in the Known Victim condition simply more confused?*

Berkowitz (1988) has suggested that "arbitrary" or unexpected costs are particularly anger-provoking. Subjects in the KV did report more surprise than subjects in the AV condition, although the difference was not significant. Could

differences in surprise be the real cause of the changes in anger? To test this possibility, a partial correlation was computed between the knowledge manipulation and the subject's change in anger, controlling for self-reported ratings of surprise. Results indicate that the effect of the offender's knowledge of victim identity still correlates with changes in anger, partial  $r = .776$ ,  $p < .001$  (simple  $r = .785$ ). Surprise was not the cause of the subjects' differential change in anger.

*Are the subjects in the KV condition more afraid of future costs?*

A central tenant of Cognitive Neoassociationism is that fear and anger can interplay such that increases in fear can cause increases in anger (Berkowitz 1990). If subjects in the KV condition believed that the offender was more likely to victimize them in the future, they could have been more fearful and this fear could have driven the change in anger. An attempt was made to minimize this possibility by specifying in the vignette that the offender frequently plays practical jokes on others. And although subjects in the KV condition did not differ significantly in fear from those in the AV condition, there was a non-significant relationship between the offender's knowledge of victim identity and fear. Those in the KV condition reported almost half a point more fear than those in the AV condition. To ensure that this difference in fear was not responsible for the increase in anger in the KV condition, a partial correlation was computed showing the relationship between the offender's knowledge of the victim's identity and the reported change in anger, controlling for fear. Results indicate that controlling for fear did not affect the relationship between the offender's knowledge of victim identity and change in anger, partial  $r = .780$ ,  $p < .001$  (simple  $r$

= .785). It is concluded that the effect of the knowledge manipulation on anger was not due to a change in fear.

*How do we know subjects are imputing a kind of intentionality rather than reacting to an unknown covariate of the knowledge manipulation?*

It is difficult to rule out unspecified and unmeasured covariates, so subjects were asked at the end of the vignette to explain why their anger changed (if it did so). Subjects' responses supported the hypothesis that those in the Known Victim condition were especially angry as a result of the offender's knowledge of their identity and those in the Anonymous Victim condition were less angry when it became known to them that the offender did not know they were the victim.

Typical examples from subjects in the KV condition whose anger increased include: "It was a malicious act, directed intently on me;" "he assumes I am passive and wouldn't do a thing about it;" "He knew what he was doing;" "he did it to me on purpose now I'm really angry 'cause now I take it personally."

Responses from the AV condition whose anger lessened also suggest that subjects were computing a kind of intentionality based on the offender's knowledge of the victim's identity. Typical examples include: "I was a random 'victim,' so it's not like the guy has a thing against me or anything;" "He didn't even know it was my lunch, so he wasn't intentionally doing it to me;" "It wasn't personal;" "It wasn't a direct prank to specifically insult me;" "I know he didn't intend to do it to me;" "My anger changed because I knew it wasn't directed at me, but my embarrassment + sadness did not change."

Responses from subjects whose anger ratings varied from these patterns were also instructive. For example, two subjects in the AV condition had anger levels that did not change and mentioned that they had assumed the offender had not known the identity of the victim before the manipulation, “[I] assumed he was a prankster and not aimed at me.” Two females in the KV condition who did not become angrier indicated that they were flattered that their lunch was chosen and suggested that he was flirting with them. Such cases, while not showing evidence in support of the hypothesis, suggest that individual exceptions may be due to (1) the subject assuming prior to the manipulation what the manipulation later confirmed, and (2) the subject not viewing the joke as having imposed a cost on them (the flirting).

**Summary and Conclusions for Experiment A2.** Actions indicating a low WTR toward the subject should trigger more anger than actions that do not. When a cost is imposed intentionally on a particular person, this reveals a low WTR toward that person. But one cannot infer that an actor has a low WTR toward a person if the actor did not know in advance who would suffer the cost of his actions. This predicts that subjects’ anger should lessen in the Anonymous Victim condition, in which the prankster did not know who would be the butt of his joke/insult, compared to the Known Victim condition, where he did. That is precisely the pattern found. The regulator theory of anger predicted this pattern in advance. This pattern cannot be explained by six of the alternative theories, and it is not predicted by the other three theories (although arguably it is consistent with those three).

***Experiment A3: Can human responses to anger be understood as attempting to indicate a high WTR and thus assuage anger?***

Evidence from the universality of the anger (see Chapter 2, feature #6) suggests that human ancestors have interacted with angry individuals for a substantial portion of human evolutionary history. Given the potentially dramatic changes in fitness that could have occurred as a result of those interactions, it is reasonable to expect that human beings are well designed to interact with angry individuals, specifically, the human mind ought to be equipped with methods designed to deactivate the anger mechanism. One way to deactivate an anger response is to target its causal inputs. If X and Y cause anger, humans are most likely designed to convince angry individuals of not-X and not-Y. In this sense an anger-deactivating mechanism in the target should fit the anger activating system in the angered person like a key fits a lock. If this is true, and it is reasonable to presume it is because of the length of selective history and the intensity of selection, one way to test a hypothesized function of anger is to look at how people respond to anger. If anger is frequently caused by a perception that someone imposed a cost that is too large given the relatively small benefit gleaned from the act, then it should be true that humans, confronted by an angry individual, should argue that the cost was in fact small, and that the benefit they received from imposing it was in fact large. If anger is exacerbated when the offender is thought to have known the identity of the person on whom the cost was imposed, then the offender should argue that they did not in fact know the identity of that person.

In Experiment A3, subjects read vignettes about situations in which the subject provoked anger in another person. Subjects were then asked which arguments they would make if they wanted to argue that what they had done was not so bad as to warrant anger. If the Regulator Theory of human anger accurately describes the causes of anger, and humans have mechanisms designed to defuse that anger, then subjects are predicted to choose arguments that emphasize that the cost imposed on the other person was small and that the benefit given to the subject as a result of imposing that cost was large. Furthermore, if anger is exacerbated when the offender knew the identity of the victim, then the subjects should argue that they did not in fact know the identity of the victim.

None of the other theories reviewed in Chapter 2 make explicit predictions (or implicit predictions that I can glean) about how individuals should respond to anger.

### **Methods for Experiment A3.**

Participants: Fifty-six undergraduates at UCSB participated for partial fulfillment of their introductory psychology class credit: 21 males (mean age 18.4, ranging 17 to 20), and 35 females (mean age 18.2, ranging 17 to 21).

Procedure. Subjects were brought into the lab in groups of two to five and were given a questionnaire that contained three scenarios. All the scenarios described the subject imposing a cost on someone else for some benefit. The order of scenarios was counterbalanced across subjects.

Materials. Subjects received instructions followed by the scenarios. The instructions and a sample scenario follow (theoretical categories are given for each argument in parentheses; these did not appear in the stimuli).

We are interested in how people make arguments. Below we will describe some situations in which someone is mad at you because of something you did. We will then list a series of statements that you could bring up in your argument. It is your job to decide which statements would strengthen your case, and which ones would weaken it.

For example, if a teacher accused you of cheating on a test, you would not defend yourself by saying, "My notes were laying open on the floor." That would weaken your case.

Several conflicts are described on the pages that follow. Read them carefully. Below each are a series of statements with a blank space before them. Some of these statements may be contradictory. **Rate each statement as if it is the only argument you are going to make on your own behalf.** Use this scale:

-3	-2	-1	0	+1	+2	+3
definitely weakens						definitely strengthens
my case						my case

- If a statement would strengthen your case, assign it a positive number. (For example, if a statement helps your case a lot, then assign it a +3; if it would only help your case a little, then assign it a +1.)
- Put a 0 next to statements that don't help or hurt your case.
- Assign negative numbers to statements that would definitely weaken your case. (For example, if a statement hurts your case a lot, then assign it a -3; if it would only hurt your case a little, then assign it a -1.)

There are no trick answers. We're interested in your gut reaction.

[page break]

Imagine that you ruined one of your friend's sweaters. It is completely useless to him now, and he is very upset about it. You want to argue that what you did was not that bad. How much would each argument strengthen or weaken your case? Assume that he will believe you when you make the statement.

- \_\_\_\_\_ His grandmother had knit him the sweater before she died. (HC: high cost)
- \_\_\_\_\_ He got the sweater at a garage sale for a dollar. (LC: low cost)
- \_\_\_\_\_ You ruined his sweater by using it to stop the bleeding from a gash on your head caused by a car crash. (HB: high benefit)
- \_\_\_\_\_ You ruined his sweater by using it to clean your tub. (LB: low benefit)
- \_\_\_\_\_ You knew it was his sweater when you ruined it. (KV: Known victim)

- \_\_\_ You didn't know it was his sweater when you ruined it. (AV: Anonymous victim)
- \_\_\_ You ruined his sweater on purpose. (distracter)
- \_\_\_ You didn't ruin his sweater on purpose. (distracter)

Another scenario involved the subject talking on the phone instead of busing tables while at work, angering a co-worker. The last scenario involved the subject stealing prescription drugs from a pharmacy, angering the owner of the pharmacy. The text of these scenarios is in Appendix A.

The subjects were given a choice of arguments that they could use if they were arguing with the person in the scenario. In particular, subjects were asked to choose statements that would help their case if they wanted to argue that what they did was not "that bad."<sup>21</sup> The order of the arguments was counter-balanced across subjects and scenarios.

The subjects were told to assume that the angry individual in each case would believe the statements they argued, so that variance stemming from the believability of the argument could be minimized. Subjects rated the arguments from -3 (definitely weakens my case) to +3 (definitely strengthens my case).

---

<sup>21</sup> This wording was chosen to prevent subjects from making arguments in the form of apologies. One can imagine at least two ways of deactivating an anger mechanism. Convincing someone they are mistaken about the input conditions is one (tested in Experiment A3). But another way of assuaging anger is to allow it to perform its function, namely to recalibrate the WTR of the offender. Apologies could be the expression of the second strategy, a way of indicating that one has in fact recalibrated one's valuation of a person or a resource in such a way as to make it no longer necessary to express anger. The content of apologies seems consistent with this function (see Chapter 2).

There were several distracter arguments into which six arguments of theoretical interest were mixed. The arguments of theoretical interest fell into the categories shown in Table 3-2.

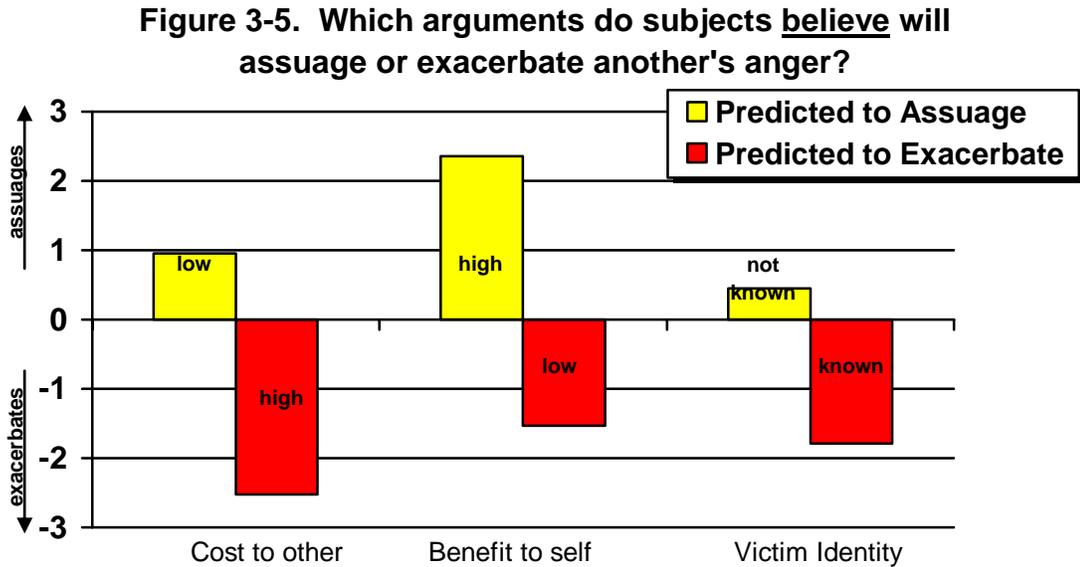
<b>Table 3-2. Predictions for Experiment A3</b>	
Arguments predicted to <b>assuage</b> anger, <b>helping</b> the subject's case	Arguments predicted to <b>exacerbate</b> anger, <b>hurting</b> the subject's case
<b>Low Cost to other (LC):</b> an argument suggesting the <u>cost</u> imposed on the other individual was <u>low</u>	<b>High Cost to other (HC):</b> an argument suggesting the <u>cost</u> imposed on the other individual was <u>high</u>
<b>High Benefit to self (HB):</b> an argument suggesting the <u>benefit</u> received by the subject was <u>high</u>	<b>Low Benefit to self (LB):</b> an argument suggesting the <u>benefit</u> received by the subject was <u>low</u>
<b>Anonymous Victim (AV):</b> an argument suggesting the subject <u>did not know</u> the victim's identity when committing the offense	<b>Known Victim (KV):</b> an argument suggesting the subject <u>knew</u> the victim's identity when committing the offense

### **Predictions of the Regulator Theory**

The predictions are detailed in Table 3-2. If anger is triggered by an implicit calculation of a Welfare Tradeoff Ratio and human beings have faced a selective history of this anger, then one would expect subjects attempting to assuage anger in another should favor arguments that imply their WTR toward the angry person is higher than that person currently believes. They should stress a low cost to the victim (LC) compared to a high cost (HC), high benefit for the subject (HB) compared to low benefit (LB), and a lack of knowledge of the victim's identity (AV) compared to knowledge of victim's identity (KV).

### **Results and Discussion for Experiment A3**

Figure 3-5 shows the average ratings subjects gave to each argument type. Average ratings over the three scenarios were computed for each subject, and the average of these taken.



The arguments predicted to “weaken the subject’s case” (i.e., to assuage anger) were all rated significantly above zero: LC ( $t(55) = 8.67, p < .00001$ ), HB ( $t(55) = 25.78, p < .00001$ ), AV ( $t(55) = 2.674, p = .01$ ). The arguments predicted to “weaken the subject’s case” (i.e., to exacerbate anger) were all rated significantly below zero: HC ( $t(55) = -20.759, p < .00001$ ), LB ( $t(55) = -12.763, p < .00001$ ), KV ( $t(55) = -9.745, p < .00001$ ).

The most critical predictions, however, concern the differences in ratings between arguments of the same category (LC>HC; HB>LB; AV>KV).

*LC > HC: When trying to assuage anger, do subjects favor arguments stressing low costs imposed over high costs imposed?*

Yes. To test the hypothesis that subjects favored low cost (LC) statements over high cost statements (HC), an average difference score was computed for each subject by subtracting their rating of the HC argument from their rating of the LC argument and averaging these three numbers (i.e. average preference for LC arguments =  $[(LC_1 - HC_1) + (LC_2 - HC_2) + (LC_3 - HC_3)] / 3$ .) This provides an indicator of how much the subject preferred LC arguments compared to HC arguments averaged across all three scenarios. This average was tested against zero with a single-sample t-test to see which arguments the subjects preferred. Results showed a large preference for the LC argument ( $M = 3.48, SD = 1.249$ ),  $t^*(55) = 20.87, p < .001, r = .95$ . On average, subjects rated the LC argument 3.48 points higher than the HC argument on the 7 point scale. In fact, 55 out of 56 subjects preferred the LC argument averaged across scenarios.<sup>22</sup>

*HB > LB: When trying to assuage anger, do subjects favor arguments stressing they gained a high benefit over those stressing they gained a low benefit?*

Yes, they judge high benefit to self arguments as more effective than low benefit to self arguments. Following the same method as above, an average difference score was computed by subtracting each subject's ratings of the LB argument from the HB argument and averaging across scenarios. When this average was tested against zero with a single-sample t-test, it showed subjects believe the HB argument will help their

---

<sup>22</sup> Research assistants reported that the single individual who did not look intoxicated.

case far more than the LB argument ( $M = 3.89$ ,  $SD = 1.08$ ;  $t(55) = 27.00$ ,  $p < .001$ ,  $r = .96$ ). No subject reported an average preference for the LB argument.

*AV > KV: When trying to assuage anger, do subjects favor arguments stressing they did not know who they were imposing a cost on over those stressing they did?*

Yes, they judged the argument that they did not know who they were imposing the cost on as more effective. Given the diversity of contexts in the scenarios, it was only possible to include statements regarding identity of the victim in the sweater-ruined scenario printed above. A single-sample t-test was conducted on the KV (Known Victim) and AV (Anonymous Victim) statements to determine if subjects preferred the AV statement. Results indicated that the anonymous victim statement was rated an average of 2.232 points higher on the 7 – point scale than the known victim statement ( $SD = 1.916$ ),  $t^*(55) = 8.716$ ,  $p < .001$ ,  $r = 0.762$ , indicating that subjects preferred to argue that they did not know on whom they were imposing a cost at the time of the transaction.

**Summary and Conclusions for Experiment A3.** Results of Experiment A3 demonstrated that subjects, when attempting to dissuade an angry individual after a provocation, will argue that the cost imposed on that person was slight, that the benefit the subject gleaned as a result of imposing that cost was large, and that they did not have knowledge as to the victim's identity. Subjects appear to be arguing as if designed to deactivate an anger mechanism that is triggered by an event that indicates a low Welfare Tradeoff Ratio (i.e. perceived to impose a large cost on the angered individual in order to get a relatively small benefit for the subject, who knew

which person would be impacted by that decision). This is consistent with the Regulator Theory of human anger and the idea that human beings have faced a long selective history of such anger and that selection has designed anger-disabling mechanisms.

**Experiment A4: Converging evidence: Which arguments in fact assuage anger?**

Experiment A4 serves as the mirror image of Experiment A3, and designed to provide converging evidence for the hypotheses about welfare tradeoff ratios explored in Experiments A1 and A2. The goal of Experiment A4 was to see whether subjects' intuitions about which arguments will serve to diffuse anger are accurate. They should be: By hypothesis, there has been selection for anger-diffusing mechanisms whose design exploits the structure of the anger-triggering mechanisms.

In Experiment A4, subjects were given the scenarios used in Experiment A3, but instead of the subject being the person who imposed the cost (as in A3), the text was rewritten so that the subject was the person *on whom the cost was imposed*. Subjects were then asked to rate which facts would have exacerbated their anger and which facts would have diminished their anger.

Experiment A1 showed evidence that subjects modulate their anger in response to the magnitude of the benefit to the offender. Experiment A2 showed that subjects report more anger when a cost is imposed on them by someone who knows their identity compared with those who have the same cost imposed on them by someone who does not know their identity. Experiment A4 is a conceptual replication of A1

and A2, but using the stimuli of A3. At the same time, it checks whether subjects' intuitions about anger-diffusion, as reflected in the arguments they think will be effective, are accurate. If they are, then the arguments thought to assuage anger in A3 should actually assuage anger in A4, and those thought to exacerbate anger in A3 should actually do so in A4.

#### **Methods for Study A4.**

Participants: Sixty-four undergraduates from UCSB participated for partial fulfillment of their introductory psychology class credit: 21 males (average age 18.4, range 17 to 20), and 43 females (average age 18.4, range 17 to 21).

Procedure. As in Experiment A3, subjects were brought into the lab in groups of two to five and were given a questionnaire that contained three scenarios based on those from Experiment A3. These scenarios were rewritten so that the subject was the person on whom the cost was imposed. Other than changes necessary to accomplish that goal, nothing was changed in the scenario. The example from Experiment A3 is shown here in rewritten form:

We are interested in how people react to arguments. Below we will describe some situations in which you are mad at someone because of something they did. They don't seem to think what they did was very bad, but you do (at least at first). We will then list a series of facts that might relate to your anger. It is your job to decide which facts would make you more angry and which facts would make you less angry. Some of these facts may be contradictory. **Rate each fact as if it is the only fact you are considering.** Use this scale:

-3	-2	-1	0	+1	+2	+3
definitely makes me						definitely makes me
less angry						more angry

- If a fact would make you more angry, assign it a positive number. (For example, if a fact makes you a lot more angry, then assign it a +3; if it would only make you a little more angry, then assign it a +1.)
- Put a 0 next to facts that don't affect your anger.

- Assign negative numbers to facts that would definitely make you less angry. (For example, if a fact makes you a lot less angry, then assign it a -3; if it would only make you a little less angry, then assign it a -1.)

There are no trick answers. We're interested in your gut reaction.

[page break]

Imagine that a friend ruined one of your sweaters. It is completely useless now, and you are very upset about it. He doesn't think what he did was very bad. How would each fact affect your anger?

- \_\_\_ Your grandmother had knit you the sweater before she died. (HC)
- \_\_\_ You got the sweater at a garage sale for a dollar. (LC)
- \_\_\_ He ruined your sweater by using it to stop the bleeding from a gash on his head caused by a car crash. (HB)
- \_\_\_ He ruined your sweater by using it to clean his tub. (LB)
- \_\_\_ He knew it was your sweater when he ruined it. (KT)
- \_\_\_ He didn't know it was your sweater when he ruined it. (AT)
- \_\_\_ He ruined your sweater on purpose.
- \_\_\_ He didn't ruin your sweater on purpose.

The second scenario involved a co-worker who was talking on the phone instead of bussing tables while the subject worked and the third scenario involved someone stealing prescription drugs from the subject's store (See Appendix A).

For each scenario, the subjects were given the same list of possible statements from Experiment A3 that could pertain to the story (the statements were rewritten when necessary to maintain the role of the subject as the person on whom the cost was imposed). As in Experiment A3, there was a statement suggesting the cost imposed was high (HC), a statement suggesting the cost was low (LC), a statement suggesting the benefit was high (HB), statement suggesting the benefit was low (LB), and, in the sweater scenario, statements suggesting the subject's identity was either

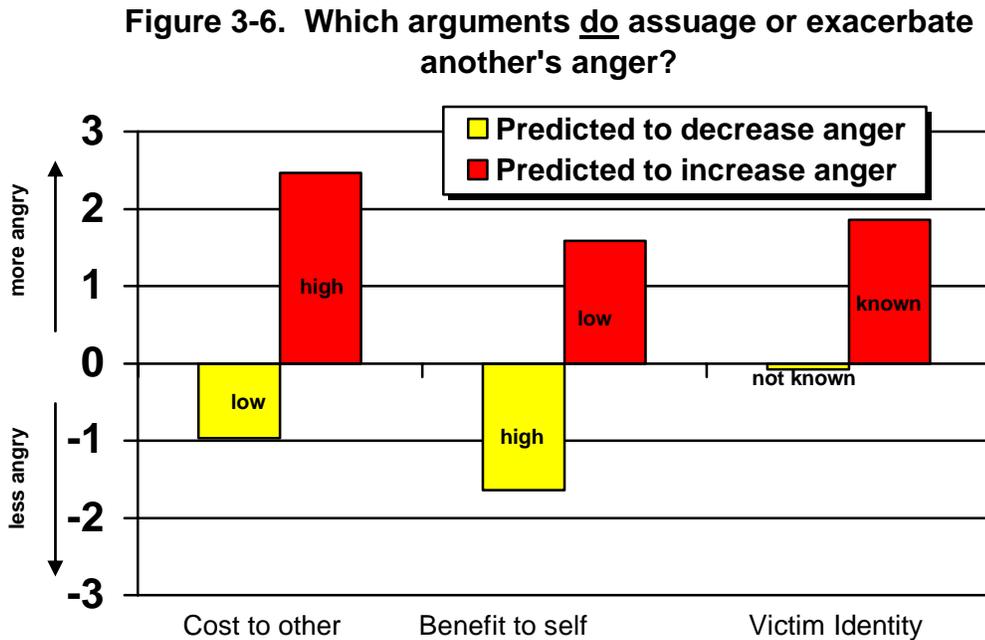
known (KV) or not known (AV). Subjects rated the statements from 1 (definitely makes me less angry) to 7 (definitely makes me more angry).

**Predictions from Regulator Theory.**

The predictions are the same as in Table 3-2. In Experiment A3, it was predicted that when subjects are in the offender role, they understand which statements will diffuse or exacerbate anger. In Experiment A4, it was predicted that those statements predicted to diffuse or exacerbate anger when subjects are in the role of offender will indeed have this effect when subjects are in the role of victim.

**Results and Discussion for Experiment A4.**

Figure 3-6 shows how subjects rated each statement.



The arguments predicted to decrease anger were all rated significantly below zero except the AV condition (anonymous victim): LC ( $t(63) = -9.107, p < .00001$ ), HB

( $t(63) = -16.323, p < .00001$ ), AV ( $t(63) = -.480, p = .633$ ). The arguments predicted to exacerbate anger were all rated significantly above zero: HC ( $t(63) = 36.074, p < .00001$ ), LB ( $t(63) = 19.389, p < .00001$ ), KV ( $t(63) = 13.059, p < .00001$ ). It is possible that the non-significance of the AV statement is because subjects assume that a friend would not intentionally ruin their sweater, and thus discovering that the friend did not know their identity would not change anger because the subjects never thought their friend did know.

*When the cost imposed is high, are subjects angrier than when the cost imposed is low?*

Yes. As in Exp A3, a difference score was computed for each subject by subtracting their rating of the low cost (LC) statement from their rating for the high cost (HC) statement on each scenario, and then averaging across scenarios. This mean difference between HC and LC statements was then tested against zero. As predicted, subjects showed greater anger in response to statements indicating the cost imposed on them was high ( $M = 3.43, SD = 0.979$ ),  $t(63) = 28.018, p < .001, r = .962$ . None of the 64 subjects showed less anger in response to a HC statement than a LC statement on any of the scenarios.

*When the offender obtains a high benefit (relative to the cost imposed) are subjects less angry than when the offender imposed the same cost to obtain a low benefit?*

Yes. A difference score was computed for each subject by subtracting their rating of the high benefit (HB) statement from their rating of the low benefit (LB) statement on each scenario and averaging across scenarios. This average was tested against

zero and showed much less anger in response to HB statements compared to LB statements ( $M = 3.22$ ,  $SD = 0.918$ ),  $t(64) = 28.09$ ,  $p < .001$ ,  $r = 0.962$ . No subject had a difference score indicating that HB statements were more anger provoking than low benefit ones.

*Is more anger elicited when the offender knew the victim's identity when imposing the cost than when the offender did not know the victim's identity?*

Yes. A paired samples t-test was run to determine if subjects reported more anger in response to the Known Victim statement (i.e., He knew it was your sweater when he ruined it) compared to the Anonymous Victim statement (i.e., He didn't know it was your sweater when he ruined it). Results replicated those of Experiment A2, with the KV statement being rated 1.938 points higher on average than the AV statement ( $SD = 1.612$ ),  $t(63) = 9.614$ ,  $p < .001$ ,  $r = 0.771$ .

*Do changes in anger in Experiment A4 track the magnitude of the preference for the argument in Experiment A3?*

A well-designed system for diffusing anger should not only be able to tell the difference between statements of high and low valuation, but should also be able to gauge the magnitude of that difference. This would not only enable the mechanism to discern the difference between statements of valuation that are closely matched, but also allow one to determine (i) when one should diffuse anger by counter-acting its causal variables, (ii) when one should allow anger to fulfill its function (recalibrate the target's WTR or estimates of costs and benefits) and apologize, and (iii) a number of other anger-reducing strategies (e.g., in the case of deception, whether it is a

reasonable strategy to lie about the valuation of a resource depends on the difference in magnitude between the real valuation and the sham valuation). Because subjects in Experiment A3 rated virtually the same statements as subjects in Experiment A4, one can correlate these statements to determine how closely the preferred arguments of subjects in Experiment A3 track the anger those statements would elicit from subjects' in Experiment A4. A well-designed mechanism for diffusing anger should produce a very high correlation.<sup>23</sup>

There were four statements regarding costs and benefits in each of the three scenarios (HC, LC, HB and LB), and two statements in one scenario regarding knowledge of the victim's identity (KV and AV), making fourteen statements of theoretical interest in both Experiment A3 and Study A4. By pairing the statements in Experiment A3 to their equivalent statements in Study A4 (e.g. "He got the sweater at a garage sale for a dollar" in Experiment A3 scenario 1 with "You got the sweater at a garage sale for a dollar" in Study A4 scenario 1) one can determine the average correlation between their use by the offender to diffuse anger and the anger they would produce. As predicted, results indicate a very strong negative correlation,  $r = -.976$ , showing that if subjects rated an argument as good for diffusing anger directed at oneself in Experiment A3, other subjects rated that same argument as effective for diffusing their own anger in Experiment A4. The ratings from Experiment A3 were reverse coded and are displayed in Figure 3-7 next to their equivalents from

---

<sup>23</sup> Mechanisms for determining the value of contested resources (both costs and benefits) would be necessary for a host of other mechanisms as well: social exchange, precautions, and decisions between alternative courses of action that lead to differential benefits and costs to various individuals. It is not my intent to imply the calculation of resource value is limited to incidents of resource competition.

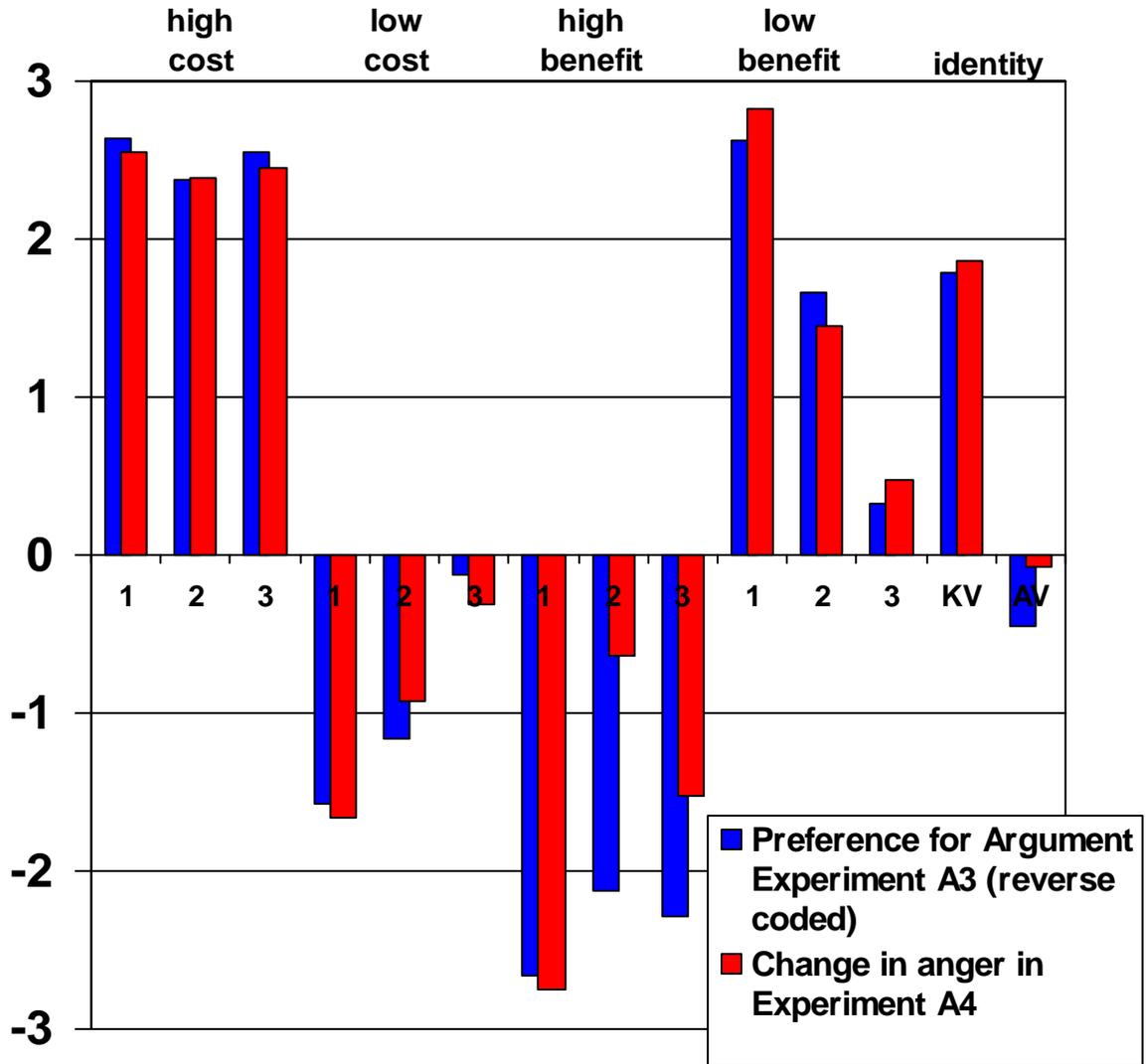
Experiment A4. Preferences for arguments by those confronted with anger appear to track almost exactly the statements that reduced anger. (There was one exception: although the values for the rating of the HB statement in the second scenario are both in the same direction, the magnitudes are not as well-correlated as for the other variables and scenarios.<sup>24</sup>) Note that the high correlation between preferences for arguments in Experiment 3A and their anger reducing effects in Experiment 4A was not due to ceiling effects that masked variation, as about half of the statements were within the middle half of the 7 – point scale (i.e. between -1.5 and 1.5).

The role of arguments in anger is further tested and explored in Chapter 5.

---

<sup>24</sup> In this scenario in Experiment A4 a woman named Sandra is supposed to be busing tables but spends her time talking on the phone while the subject must work her shift. The HB statement is: Sandra is talking to her boyfriend and finally convinces him to marry her during that phone call. Six subjects rated this as highly anger provoking (rating it as a 2 or 3 on a scale from -3 to 3). One can only speculate on the reason for these anomalous scores. A colleague suggested that male subjects may be more willing to be angry with a female coworker who is engaged because she is not available for a romantic encounter, however, male ratings of the HB statement averaged -1.14 while female ratings averaged -.39. Furthermore, only one of the 6 who rated the statement as highly anger provoking was male. A research assistant unfamiliar with the hypotheses of the study suggested that some female subjects might think it was “stupid to get engaged over the phone.” This could be interpreted as indicating that the benefit of getting engaged over the phone is lower than intended. Subjects could also have interpreted the need to “convince” her boyfriend to get engaged as indicating low status in Sandra (suggesting she is to some extent pathetic if she must beg her boyfriend to marry her).

Figure 3-7: Preference for argument in Experiment A3 (reverse coded) with the argument's impact on anger in Experiment A4.



*Do the results of Experiments A3 and A4 generalize outside the laboratory?*

Using vignettes allows one to ask questions that would be unethical to address outside the laboratory, but there is always the possibility that the results gleaned from such tasks would be different if they were run outside the laboratory. One way of testing this possibility is to include questions that test effects known to be present in more ecologically valid experimental designs.

Four statements were included in the distracter items of Experiments A3 and A4 that test the effect of general intentionality on anger: a relationship that has been replicated across a number of experimental designs. Though researchers have had a hard time making a computational model of exactly what intentionality is, subjects know roughly what it means when something is done “on purpose.” Two statements indicated that the cost was imposed either “on purpose” or “not on purpose.” Also, two statements were added indicating one either did or did not have knowledge of the cost being imposed. Subjects rated how likely they would be to use these statements in Experiment A3, and how likely they would be to cause anger in Experiment A4. See Table 3-3 for the versions of those statements that the female subjects received.

**Table 3-3: Statements of Intentionality in Experiments A3 and A4.**

	High Intentionality	Low Intentionality
Experiment A3 – Scenario 1	You ruined her sweater on purpose.	You didn’t ruin her sweater on purpose.
Experiment A3 – Scenario 2	You knew how long you were talking.	You didn’t realize how long you were talking.
Experiment A4 – Scenario 1	She ruined your sweater on purpose.	She didn’t ruin your sweater on purpose.
Experiment A4 – Scenario 2	Sandra knew how long she was talking.	Sandra didn’t realize how long she was talking.

Note that Scenario 2 described a person (the subject in Experiment A3) talking on the phone instead of doing work. Scenario 1 described someone (the subject in Experiment A3) ruining a sweater.

If the scenarios in Experiment A4 were in fact engaging anger mechanisms in a way similar to laboratory studies with electric shocks, interference, aversive noise, and insulting evaluations, then one would expect subjects in that experiment to rate statements indicating high intentionality as more anger provoking. They did: the average high intentionality rating was  $M = 2.406$ ,  $SD = .51$ ; the average low intentionality rating was  $M = -.586$ ,  $SD = .136$ . For each subject, their rating of high intentionality was averaged and their ratings of low intentionality were averaged, and these averages were tested against each other with a paired samples t-test. Results showed that subjects reported more anger in response to the high intentionality statement than the low intentionality statement (mean difference = 2.99,  $SD = 1.14$ ),  $t(63) = 21.01$ ,  $p < .0001$ .

Subjects in Experiment A3 chose arguments that emphasize lack of intentionality: low intentionality  $M = 1.223$ ,  $SD = .774$ ; high intentionality  $M = -2.080$ ,  $SD = .96$ ). The difference between the subjects' average ratings of high intentionality and average ratings of low intentionality were tested with a paired samples t-test. Results showed a strong preference for low intentionality statements on the part of those who had imposed a cost, ( $M = -3.306$ ,  $SD = 1.267$ ),  $t(55) = -19.507$ ,  $p < .00001$ ,  $r = .935$ .

As with the cost, benefit, and identity variables, the intentionality statements that subjects in Experiment A3 preferred when they had imposed a cost were the same

statements that lessened anger in subjects in Experiment A4. When the average ratings of the four statements of high or low intentionality from Experiment A3 were correlated with those from Study A4 there was a very strong correlation,  $r = -.980$ ,  $p < .05$ . Again, this is not the result of a ceiling effect, as only two of the eight statements averaged above 2 or below -2 on a scale that ranged from 3 to -3. Thus, once again, subjects who have imposed costs on others are choosing arguments that would reduce anger. The fact that the results for variables that have been validated with behavioral measures are giving the same results with the vignette studies supports the view that the vignette studies with self-reported anger are tapping the same mechanisms.

One more point should be made. It is not clear that behavioral measures of anger, such as willingness to shock another person, should be preferred to self-report measures. Prudential factors can and should affect one's willingness to exact punishment, but that is not true for a self-report of anger. Secondly, to the extent that anger reflects the activation of a mechanism, of which there is an internal signal indicating degree of activation, self-report of the internal experience may be the most straightforward and valid measure.

**Summary and Conclusions for Experiment A4.** The results of Experiment A4 replicated the results of Experiment A1 and A2. Across three different scenarios involving a merchant, co-worker, and friend, subjects were angrier when a large cost was imposed on them and when a relatively small benefit was gleaned by the offender as a result of imposing that cost. In the only scenario in which it was tested, subjects

were more angry when the offender knew their identity at the time they imposed the cost.

### *E. Discussion of Experiment Set A*

Experiments A1-A4 demonstrated that anger is reliably evoked when individuals impose a cost-benefit transaction that indicates the use of a low Welfare Tradeoff Ratio toward the individual who becomes angry. All else equal, more anger is elicited when: i) the cost imposed is large (see Chapter 2 feature #1, ii) the benefit gained by imposing that cost is low (see Experiment A1 and A4), iii) and the individual who imposed the cost was aware of the magnitude of the cost-benefit transaction and the identity of the individual on whom the cost was imposed (see Experiment A2 and A4). These findings confirm the predictions of the Regulator Theory. In chapter 6, I will consider their impact on alternative theories.

## Chapter 4.

### Individual Differences in Male Anger – Study Set B.

Studies of individual differences in anger have been more common than studies of the normal functioning of anger. This appears to be because of the enormous applied value of such information in detecting and preventing the aggression that sometimes results from anger. Over the past thirty years research has linked all of the following to increases in aggression or anger: alcoholism, violent media consumption, heritability, testosterone and a host of personality variables such as narcissism, self-esteem and emotional instability (see Geen, 1998). The effects of many of these variables are undoubtedly by-products (i.e. effects on the phenotype that were not functionally designed by natural selection). For example, given that alcohol in quantities sufficient to create drunkenness was unavailable during most of human evolutionary history (human production of alcohol through fermenting requires bowls, which are relatively recent), one cannot posit that natural selection designed humans to respond to alcohol with increased aggression. Other predictors of individual differences in anger such as narcissism or high self-esteem fit nicely into the Regulator Theory of Anger.<sup>25</sup> One aspect of the phenotype that is predicted by the Regulator Theory of human anger to influence thresholds of anger has been neglected, however: physical strength in human males.

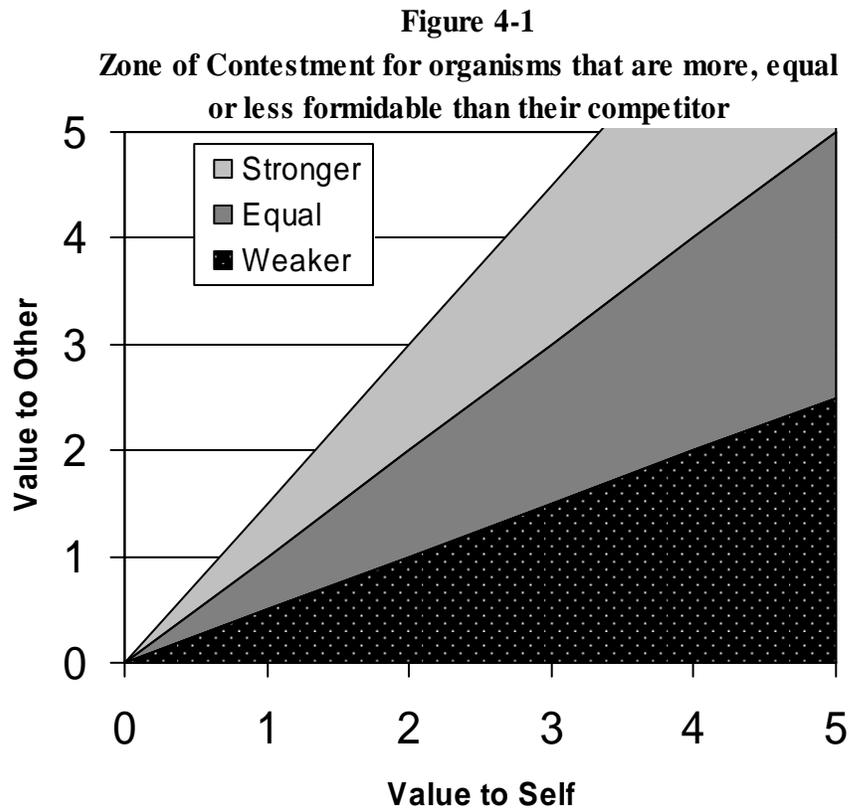
#### *A. Formidability and Human Anger*

##### **1. Formidability predicts aggression in non-humans**

Imagine two organisms. Now, imagine the set of all possible resource competitions with asymmetric value to these two contestants (e.g. conflict over a piece of food worth 3 to one contestant and 6 to another; conflict over a mate worth 2 to one contestant and 12 to the other; etc.) According to the Asymmetric War of Attrition, for each individual in a pair, there will be a subset of conflicts of interest that organisms will be selected to relinquish without dispute (those with a high value to the other and a low value to one's self), and a distinct subset that organisms will be selected to contest (those with a low value to the other and a high value to one's self). The size of that zone of contestment will be determined by the relative Resource Holding Potential of the two organisms. The more formidable the organism relative to its competitor, the larger the zone of contestment will be, and the fewer resources will be relinquished without a fight (see Figure 4-1).

---

<sup>25</sup> This assumes, of course, that these personality traits reference the expectation of higher WTRs directed at one's self from others, as seems likely.



Note: Stronger organisms will be selected to contest all resources in and under the “stronger” area; an organism facing an equally matched competitor will contest in and under the “equal” area; and weaker organisms will be selected to only contest those resources in the “weaker” area.

For many organisms, size predicts RHP, such that larger individuals are stronger and therefore more able to win aggressive contests. When this is true, these selection pressures predict that, other factors equal, larger (i.e. more formidable) organisms will be more likely to initiate combat over contested resources and less likely to relinquish them.

There is abundant evidence suggesting that larger organisms are more aggressive across a broad spectrum of non-human animals including sea anemone (Brace &

Pavey, 1978), mollusks (Zack, 1975), hermit crabs (Mitchell, 1976; Dowds & Elwood 1983; Neil, 1985), field crickets (Dixon & Cade, 1986), bluegill sunfish (Henderson & Chiszar, 1977), brown anoles (Tokarz, 1985), rats (Robitaille & Bovet, 1976), and a host of primates (Smuts, Cheney, Seyfarth, Wrangham & Struhsaker, 1987).<sup>26</sup>

## **2. Formidability should predict aggression in human males**

Imagine two humans. Now, imagine the set of all possible conflicts of interest with asymmetric value to these two individuals. One's Welfare Tradeoff Ratio will determine the subset of those conflicts for which one will relinquish and the subset for which one will contest. To the extent that physical aggression was used to enforce Welfare Tradeoff Ratios in the environment in which anger evolved, the Regulator Theory of Human Anger would predict that physically stronger individuals, like other animals, would be designed to force others to use higher WTRs toward them, and thus would respond with anger over a greater range of possible conflicts of interest. Given the ubiquity of violent human conflict among males (Daly & Wilson, 1988; LeBlanc 2003), it seems certain that physical strength played *some* role in outcomes of resource competition for men, and thus one would predict that stronger males would be more prone to anger. That is, all else equal, the same cost-benefit transaction should be more likely to elicit anger in stronger than in weaker men.

Previous work by criminologists using Sheldon's somatotypes classification of male physique (see Sheldon, Dupertius & McDermott, 1954) have consistently shown that mesomorphic (i.e. muscular) individuals are overrepresented in criminal and

---

<sup>26</sup> John Archer deserves credit for gathering most of this dataset. See Archer (1988) chapter nine for

delinquent populations (Sheldon, Hartl & McDermott, 1949; Epps & Parnell, 1952; Glueck & Glueck, 1956; Gibbens, 1963; Shasby and Kingsley, 1978; Cortes and Gatti, 1972). Such findings, though lending credibility to the idea that attitudes and aggression can be correlated with physique, are difficult to interpret and do not speak to the point of causation. It may be that jurors or custodians are more likely to convict muscular men or that being in prison leads men to develop muscular bodies. It further remains untested whether muscular body types are in fact more physically strong (as they tended to be shorter than ectomorphs and weigh less than endomorphs). It is important to distinguish the result—that mesomorphy correlates with being incarcerated or delinquent—from Sheldon’s theory of body types, which was based on an outdated model of embryology that explained relationships between aggression and body type as an accident of development.

**3. More formidable males should view aggression as a more effective way to resolve interpersonal disputes and be less afraid of retaliation.** The stronger a man is, the greater the WTR others should hold toward him. This means that others will be less likely to retaliate against him when he uses aggression in resource conflicts. Because weaker males are predicted to back down in the face of stronger ones, stronger males should view the use of force as more effective than weaker males should.

**4. More formidable males should view aggression as a more effective way to resolve international disputes and be less afraid of retaliation in such disputes.**

---

*many* more examples. See also Huntingford and Turner, 1987.

This is a byproduct hypothesis, stemming from the fact that anger evolved in an ancestral social world that was small in scale, lacking geopolitics. Modern studies of foraging societies (Boehm, 1984; Chagnon, 1983; Daly & Wilson, 1988; Sargent, 1974) and archeological sites of prehistoric man (Keeley, 1996; LeBlanc, 2003) suggest that ancestral warfare most likely consisted of raiding rather than full-scale warfare. When engaged in coalitional aggression with a small number of individuals—group fights and raids can be 2-4 men—one’s personal formidability contributes to the formidability of the group. If so, then one’s own aggressive formidability would be correlated with the aggressive formidability of coalitions in which one participates.

The mind did not evolve in a world with nation states. To the extent that modern political attitudes are the result of a mind designed to see the world in terms of small groups and individual action, individual strength may influence – as a byproduct of how minds are organized – attitudes about war, the effectiveness of using force in international (as distinct from *interpersonal*) disputes, and the fear of retaliation when force is used in such disputes.

Study Set B explores new methods of measuring strength and aggression and tests novel predictions derived from the Regulator Theory of Human Anger regarding individual differences in anger and anger-related attitudes against alternative theories of anger.

***B. Predictions regarding strength and anger from other theories***

Note that most views of anger do not distinguish rumination from other forms of anger and aggression. Therefore predictions for anger will be assumed to generalize to rumination, unless the theory specifies otherwise.

### 1. Cognitive-Motivational-Relational Theory

This theory makes no explicit predictions about size and aggression, but does hold that individuals respond to cost-benefit calculations in anger. Such calculations could result in physically stronger males being more angry and aggressive because of the presumed lower costs of fighting for them.

*Prediction: Stronger males should be more angry and aggressive in interpersonal disputes, but not for political aggression.*

### 2. Cognitive Neoassociationism and Frustration-Aggression Model

The Frustration-Aggression Model and Cognitive Neoassociationism predict that individuals who are more frequently frustrated will be the ones who exhibit aggression the most frequently. Those who are less commonly frustrated will in fact be less likely to exhibit aggression.

To the extent that strength is relevant to solving conflicts of interest for males in the modern world, the Frustration-Aggression Model and Cognitive Neoassociationism predict that stronger males will be more likely to be deferred to in resource conflicts, more likely to win the disputed resource if conflict ensues, and thus be *less* likely to be frustrated. As a result of this, they should be *less* aggressive, and presumably, less angry in general.

On the other hand, Berkowitz's theory posits that, "any kind of negative affect, sadness as well as depression and agitated irritability, will produce aggressive inclinations and the primitive experience of anger before the higher order processing goes into operation" (Berkowitz, 1989, p. 71). Thus, frustration leads to anger but that anger may or may not lead to aggression depending on prudential factors - "the strength of an individual's aggressive tendencies is directly associated with the extent that he anticipates punishment or disapproval for aggression [1962, p. 93]."

Assuming the costs of punishment are less for stronger males, one would predict they would be less likely to experience anger, but, given an instance of anger, stronger individuals would be more likely to translate that anger into physical aggression.

Thus in total, Cognitive Neoassociationism could make any prediction about the relationship between strength and physical aggression, but clearly predicts that strong individuals, to the extent that they experience less negative affect, will be less angry.

Both theories predict no relationship between strength and political aggression.

*Prediction: Frustration-Aggression predicts stronger males should be less angry; Cognitive Neoassociationism is ambiguous with respect to physical aggression, but predicts stronger individuals will experience less anger. Both predict no relationship between strength and aggression.*

### 3. Equity Theory

Donnerstein and Hatfield (1982) did not extract predictions about individual differences in human anger. But let us add an assumption based on the same prudential argument above: that stronger males are more likely to be deferred to and

therefore get what they want. Equity theory clearly predicts that those who benefit more, in general, will have less reason to believe they have been treated inequitably, and thus be less likely to experience anger. With the prudential assumption added, equity theory would predict that stronger males should experience less anger.

*Prediction: None. With additional assumptions: Stronger males should be less angry*

#### 4. Behaviorism

According to a behaviorist analysis, individual differences in anger and aggression stem from a pattern of differential reinforcement, with those individuals who benefit from their anger and aggression becoming more angry and aggressive over time. To the extent that anger and aggression was more likely to result in positive outcomes for stronger individuals over their ontogenetic development, then those individuals should be more prone to anger and aggression.

It should be noted, however, that this increase in anger and aggression should be limited to domains of aggression and anger in which individuals could have had reinforcement (e.g. personal aggression, but not political aggression). The possibility of attitudes about personal aggression generalizing to political aggression is difficult to predict as behaviorism does not make clear which features are used to generalize from one class of stimuli to another. Skinner emphasized “simple physical dimensions” (Herrnstein, 1978). On these grounds, there is no basis for generalizing reinforcement in interpersonal situations to hypothetical geopolitical domains.

*Prediction: Stronger males should be more angry and aggressive, but their political attitudes should not be influenced by their personal strength.*

#### 5. Social Learning Theory

Individual differences in aggression, according to social learning theory, are the result of differential reinforcement (as with behaviorism) and also differential vicarious reinforcement, such as that observed in the media. It is not clear that the vicarious reinforcement of strong males differs from that of weak males, but the direct reinforcement is clear, as with Behaviorism.

*Prediction: Stronger males should be more angry and aggressive, but their political attitudes should not be influenced by their personal strength.*

#### 6. Social Constructivism

It is difficult to derive predictions from Social Constructivism; the only guideline is that social behavior and attitudes should be organized in ways that benefit “society.” It is not exactly clear what that means (see Chapter 2), but one reasonable prediction from it is that more powerful people should be more restrained. Society benefits when it’s most dangerous individuals demonstrate restraint – note that firearms require registration for the benefit of society because of their ability to do damage. Guns are designed with numerous safeguards precisely because, according to a social constructivist position, society needs to constraint violence. If the minds of males are also products of society, one would predict that the most dangerous of those men would be designed to minimize violent expression.

*Prediction: Stronger males should be less angry and aggressive*

## 7. Social Interactionism

If stronger males are more likely to win conflicts over resources (as seems likely), then stronger males should, according to Social Interactionism, be more prone to anger and aggression. Importantly, this should only apply to conflicts over which personal formidability is relevant. Unlike behaviorism and social learning, Social Interactionism clearly predicts that physical strength should not be a factor in determining one's attitudes about political aggression. Instead, attitudes about political aggression should be based on one's probability of paying the costs of warfare (i.e. probably of being sent to war).

*Prediction: Stronger males should be more angry and aggressive in interpersonal aggression but not for political aggression.*

## 8. "Short Man Syndrome" or "Napoleon Complex"

Though not explicitly advanced in the scientific literature on emotions or aggression, a common perception among both academics and laymen is that shorter males will overcompensate for their lack of stature by increasing aggression. Though often described in terms of height, the same argument would predict that weaker males would need to compensate for their lack of strength by increasing their proneness to anger (i.e. "the smaller the dog the louder the bark").

*Prediction: Stronger males should be less angry*

### ***C. Study B1: Strength predicts anger in college weight-lifters***

The regulator theory predicts that stronger males will, as a result of a higher expected WTR, be more prone to anger and physical confrontations, be more likely to

view physical aggression as an efficacious way of resolving interpersonal conflicts, and be more likely to have this attitude spill over into the geopolitical realm. To test these predictions, a study with sixty-three UCSB students who frequented the campus gym was conducted.

The following primary hypotheses were tested:

1. Stronger males should be more prone to anger.
2. Stronger males should NOT be more prone to angry rumination.
3. Stronger males should have been in more physical fights.
4. Stronger males should believe more in the utility of aggression in resolving personal conflicts (which entails being less likely to fear retaliation when aggression is used).
5. Stronger males should believe more in the utility of aggression in resolving international conflicts.

### **Methods for Study B1.**

Participants. Sixty-three males (average age: 21) were recruited with fliers and posters outside the Recreation Center. They were told they would participate in a study during which they would fill out a questionnaire and indicate strength by working out on a series of machines. One subject was removed from analysis because he had a physical disability that complicated the measure of his strength (he had lost a leg).

Procedure. Subjects were run one at a time through a four part procedure.

- 1). The subject was given a questionnaire containing scales designed to measure *Proneness to Anger, Tendency to Ruminare, History of Fighting, Utility of Personal Aggression* and *Utility of Political Aggression*.
- 2). When the questionnaire was completed, each subject was given a pair of black sweatpant shorts to change into. After changing into the shorts and their lifting t-shirt in the Recreation Center's men's room, a series of morphological measurements were taken by the experimenter, including height, weight, hip circumference, waist circumference, chest circumference, unflexed bicep, flexed bicep, and neck circumference.
- 3). After the body measurements, the subjects removed their t-shirts and were photographed from four perspectives: facing front full body, facing side full body, facing back full body, and facing forward face only.
- 4). Subjects were taken into the weight lifting area. They were then lead by the experimenter through five exercises in random order: arm curl (biceps), abdominal crunch (abdominal muscles, e.g. rectus abdominus), leg press (quadriceps), chest press (pectorals), and super long pull (latissimus dorsi – "lats", deltoids). For each weight lifting exercise, the subject was asked to estimate his maximum 10-repetition lifting strength on that exercise. Then he was asked to lift 10 repetitions of the following weights:

50% of maximum estimated weight

75% of maximum estimated weight

100% of maximum estimated weight

maximum estimated weight + 5 pounds

maximum estimated weight +10 pounds

maximum estimated weight +15 pounds... etc.

The most weight subjects could successfully lift ten times was recorded for each of the five exercises. These scores were then converted to z-scores and averaged together to create a single composite z-score representing the strength of each subject.

### **Results of Study B1**

*Scale Validation.* In order to establish the inter-item reliability of the scales, the items were subjected to an inter-item reliability analysis. When deletion of an item would improve the reliability of the scale by more than .01 then the item was removed from the scale.

Unless otherwise indicated, instructions for all scales read:

*Please use the following scale to answer the questions below*

<i>strongly</i>							<i>strongly</i>
<i>disagree</i>							<i>agree</i>
1	2	3	4	5	6	7	

Subjects then read items and circled the number that best reflected their agreement or disagreement with the item. Items from different scales were not kept together.

Items that were reverse coded are marked with (R).

Scale #1: Proneness to Anger: The proneness to anger scale originally consisted of the following 12 items:

1. Although I don't necessarily act on it, I feel an urge to punch people who think they are better than me.

2. People who get in my face bug the hell out of me.
3. It really bothers me if someone has gotten away with something at my expense.
4. Rate how much of a temper you have (compared to other male UCSB students).  
[Scaled from 1 (much less) to 7 (much more)]
5. If someone insults me I just let it pass. (R)
6. If another driver cuts me off, I do not get angry. (R)
7. It is harder to get me angry than other people. (R)
8. Some people just need to be taken down a peg or two.
9. If someone shoves me I shove back.
10. If someone was making too much noise in a movie theater, and ruining it for the rest of us, I would tell the loudmouth to shut up.
11. I don't back down.
12. Other people's status games don't bother me. (R)

Inter-item reliability was improved by the removal of item #12. The final Cronbach's alpha for the proneness to anger scale was  $r = .70$ .

Scale #2: Tendency to Ruminare: The tendency to ruminare scale consisted of the following 5 items:

1. I live by the motto "let bygones be bygones." (R)
2. Rate how long you hold a grudge (compared to other male UCSB students).  
[Scaled from 1 (much less) to 7 (much more)]
3. Sometimes I stay mad for days.
4. I get over being angry in an hour or two. (R)

5. Rate how irritable you are (compared to other male UCSB students). [Scaled from 1 (much less) to 7 (much more)]

The Cronbach's alpha for the tendency to ruminate scale was  $r = .78$ .

Scale #3: History of Fighting: The subjects' past experience with physical aggression was measured by the following 6 items:

1. I used to fight a lot as a kid.
2. I have gotten so angry that I threw myself at someone without thinking of the consequences.
3. I have physically intimidated someone who had it coming.
4. I have physically defended myself against attack.
5. I have stared people down.
6. Since the age of 14, I have been in fights.

Removing item #1 improved the inter-item reliability resulting in a final Cronbach's alpha for the History of Fighting scale of  $r = .82$ .

Scale #4: Utility of Personal Aggression: The goal of this scale was to assess whether individuals view aggression as an effective tool in conflicts, likely to minimize future retaliation (rather than inflame it). This scale originally consisted of the following 9 items:

1. If I don't respond to provocations and do something to make wrong-doers pay, they'll just do more to hurt me in the future.
2. If someone gets out of line with me, I think it is better to let it pass. (R)

3. If someone hurts me, and I do something back to make them pay, they'll just do more against me. (R)
4. It's not worth my time or effort to pay back someone who has wronged me. (R)
5. If a teaching assistant has been unfair to me, I don't say anything - arguing with him would just make him treat me worse in the future. (R)
6. I don't tell my father when he is being unfair to me, because he'll just get angrier. (R)
7. If someone does something to hurt me, and I don't get them back, then they'll think they can do whatever they want to me.
8. Sometimes, you just have to settle things with physical force.
9. A wise person avoids competition. (R)

Reliability was improved with the deletion of items 5 and 6 from the scale, leaving a 7 item scale with a Cronbach's alpha of  $r = .63$ .

Scale #5: Utility of Political Aggression: This scale originally consisted of the following 7 items:

1. Violence only breeds more violence. (R)
2. It is usually better to show mercy than to punish the violent. (R)
3. When used against guilty murderers, the death penalty is a well justified part of the criminal justice system.
4. I think the US spends too much on the military. (R)
5. I think it was good for the US to show the world, by going to war against Iraq, that it has the strongest military in the world.

6. I agreed more or less with the campus antiwar demonstrations. (R)

7. I think it is a mistake to go after Middle Eastern countries that harbor terrorists.

This will just lead to more attacks on America in the future. (R)

Reliability was improved with the deletion of item 2, leaving a 6 item scale with Cronbach's alpha of  $r = .85$ .

*Lifting Strength:* The heaviest weight lifted for 10 repetitions was used as the subject's score for that weight lifting task. These scores were converted to z-scores and then averaged across the five weight lifting exercises for a single composite lifting strength z-score.

#### *Other strength measures*

In an attempt to find measures of physical strength that could be used in the absence of weight-lifting machines, three predictors of strength were evaluated. These were: morphological predictors of lifting strength, self-report, and ratings of the individual's physical strength based on photographs.

#### 1. Morphological predictors of lifting strength

In order to determine the best morphological predictors of strength, a series of body measurements were taken and correlated with actual lifting strength. Their simple Pearson correlations are shown in Table 4-1.

**Table 4-1: Morphological Correlates of Lifting Strength**

	Height	Weight	Hip Circum.	Waist Circum.	Chest Circum.	Neck Circum.	Flexed Bicep Circum.	Unflex. Bicep Circum.
Lifting Strength	.255	.507	.073	.099	.589	.364	.735	.685

-Pearson Product Coefficients,  $n = 62$

In order to detect interactions, a linear regression analysis was done predicting lifting strength with all morphological variables simultaneously entered. Body measurements predicted 63.9% of the variance in lifting strength (adjusted R-squared = .584), with flexed bicep circumference being the only variable that accounted for unique variance (standardized Beta = .708,  $p=0.002$ ). Waist circumference was a marginally significant predictor of lifting strength after controlling for flexed bicep circumference (partial  $r = -.209$ ,  $p=.13$ ).

In this population, the best single predictor of lifting strength was flexed bicep circumference, and the best 2-factor predictor of lifting strength was flexed bicep circumference controlling for waist circumference.

## 2. Self-report of physical strength

A single question was inserted into the questionnaire to determine if subjects could accurately predict their strength relative to peers. The subjects were instructed to fill in the blank in the following statement: "I am stronger than \_\_\_\_\_% of other men." The correlation between their estimates and their actual lifting strength was .643,  $p<.0001$ . Furthermore, the partial correlation between lifting strength and the subjects' report of their own strength while controlling for flexed bicep circumference and photo ratings of strength (see below) was still significant (partial  $r = .255$ ,  $p=.05$ ), suggesting that subjects' self-report accounted for variance that was not predicted by flexed bicep circumference and photo ratings of strength.

## 3. Ratings of the subjects strength based on photographs of their bodies

Full body photographs of the subjects wearing a standard black pair of shorts and standing next to an experimenter (for standard comparison) were taken and edited with Photoshop 8.0 so that the subject's head was covered. These full-body photographs were then shown to a sample of 50 UCSB undergraduates (18 male and 32 female). The undergraduates rated the photographs on physical strength using a 7-point Likert scale rating from very weak (1) to very strong (7). Photographs were presented in random order.

The results were averaged giving each subject in Experiment B1 a single score indicating their strength as perceived by the raters. The correlation between actual lifting strength and the perceived strength of the weight-lifters was .623,  $p < .0001$ . Partial correlations revealed that the ratings of strength continued to predict lifting strength even after controlling for flexed bicep circumference and self-reported strength, partial  $r = .330$ ,  $p = .011$ .

#### 4. Composite Measure of Proxies of Strength

Because flexed bicep circumference, self-reported strength, and ratings of strength based on the subjects' photographs all independently predicted lifting strength, a composite measure of these variables was created. Each variable was converted into a z-score and averaged to arrive at a single composite measure of the proxies of strength. This measure correlated more highly with lifting strength than any single predictor,  $r = .794$ .

*Were stronger males more prone to anger,? To fighting? Were they more likely to believe aggression is efficacious?*

Yes, they were. Table 4-2 shows the correlation matrix with three measures of physical strength predicting subjects' scores on the Proneness to Anger, Rumination, History of Fighting, Utility of Personal Aggression, and Utility of Political Aggression Scales. The p-values are one tailed, because they involve prior predictions.

**Table 4-2: Correlations between Physical Strength and Attitudes on Anger and Aggression.**

	Proneness to Anger	History of Fighting	Utility of Personal Aggression	Utility of Political Aggression	Rumination
<b>Lifting Strength</b>	.356 p = .002	.455 p = .0001	.331 p = .004	.271 p = .017	.017 p = .448
Flexed Bicep Circumference	.451 p = .0001	.411 p = .0004	.342 p = .003	.298 p = .009	.080 p = .269
"I am stronger than <blank>% of other men.	.330 p = .004	.372 p = .001	.232 p = .035	.236 p = .033	.071 p = .292
Photo Ratings of Strength	.405 p = .001	.388 p = .001	.294 p = .011	.261 p = .021	.191 p = .07
Composite Measure of Strength	.470 p = .00006	.467 p = .0001	.344 p = .003	.309 p = .007	.139 p = .141

n = 62, one-tailed (all were directional predictions except for rumination, where p values should be doubled)

Hypothesis 1: Proneness to anger

It was predicted that physically strong males expect others to value their welfare more than weak males (i.e. set a larger zone of contestment), and as a result of this be generally more likely to experience and express anger. By all five measures of physical strength (lifting strength, bicep, self-report, photo rating and the composite measure) stronger individuals reported being more prone to anger than weaker ones.

### Hypothesis 2: Rumination

It was predicted that physically strong males would not be more likely to ruminate over a given instance of anger, as rumination is theorized to be a process used by individuals who cannot engage in confrontation to recalibrate the Welfare Tradeoff Ratio of the offending individual. Weaker males would have had to rely more on social support, opportunity, and avoidance than physically stronger males, and thus would be more likely to ruminate over offenses. On the other hand, prudential reasons—especially when there are police, courts, and lawyers—can also prevent a stronger male from engaging in confrontation. Moreover, if stronger males anger more easily, they would more often be in a position to have something to ruminate about (even if they confronted more often than weaker men). So no correlation, negative or positive, was predicted between strength and rumination.

There was no significant correlation between rumination and any of the measures of physical strength (there was, however, a marginal, one-tailed, (positive) correlation with the photograph ratings of strength).

### Hypothesis 3: History of Fighting

It was predicted that physically stronger males would report having been in more physical fights as a result of their increased proneness to anger. This hypothesis was confirmed; scores on the History of Fighting Scale correlated significantly with all measures of physical strength.

### Hypothesis 4: Utility of Personal Aggression

Attitudes about the efficacy of aggression were hypothesized to be correlated with physical strength. This hypothesis was confirmed; scores on the Utility of Personal Aggression correlated significantly with all measures of physical strength.

#### Hypothesis 5: Utility of Political Aggression

Attitudes about the efficacy of political aggression (i.e. warfare) were hypothesized to be correlated with physical strength. This hypothesis was confirmed; scores on the Utility of Political Aggression correlated significantly with all measures of physical strength.

*Is there reason to believe that stronger males are winning more conflicts of interest?*

A number of theoretical positions make predictions about individual differences in aggression based on the assumption that stronger males are more likely to win conflicts of interest, and thus experience less inequity, a shorter history of negative reinforcement for aggression, or other such effects. This assumption is difficult to verify empirically, however there is some evidence from this data set suggesting that, in at least one domain known to be important for college-aged males, stronger males are more likely to win conflicts of interest, namely sexual access.

The correlation between lifting strength and the age at which the subject first had sexual intercourse (as indicated by self-report on the questionnaire) was  $r = -.349$ ,  $p = .01$ , two-tailed. Furthermore, the questionnaire asked subjects if they were in a committed romantic relationship; results of an independent samples t-test showed that those in a committed relationship were slightly stronger than those without a romantic partner,  $t(58) = 1.192$ ,  $p > .10$ . These results show some supporting evidence for the

assumption that physically stronger males are more likely to win conflicts of interest against weaker males, at least in the sexual domains.

### **Discussion for Study B1**

Strong support was found for all five hypotheses. The lifting strength of the college students attending the Recreational Center was correlated with their proneness to anger, history of fighting, and beliefs about both personal and political aggression. As predicted, lifting strength was not correlated with angry rumination; it is noteworthy that rumination dissociated from other forms of anger, as predicted.

Two questions are addressed below: 1) can alternative theories of anger and aggression account for these findings, and 2) what non-theoretical objections can be raised about this dataset.

#### 1. Can alternative theories of anger account for these effects?

Each theory makes a different set of predictions about whether strength is correlated with each scale, and if so, whether that correlation should be positive or negative. These predictions are summarized in Table 4-3.

**Table 4-3: Summary of Predictions and Results from Study B1**

	Prone to anger	Fighting history	Utility of Personal Aggression	Utility of Political aggression	Rumin-ation	Total score
<b>Results</b>	+	+	+	+	<b>0</b>	
Regulator theory	+	+	+	+	0	5
Cog.-Mot.-Relat.	+	+	+	0	0	4
Frustration-aggression	-	-	-	0	-	-3
Cognitive-neoassoc.	-	+	+	0	-	1
Equity theory	-	-	0	0	-	-2
Behaviorist theory	+	+	+	0	0	4
Social learning theory	+	+	+	0	0	4
Social constructivism	-	-	-	0	-	-3
Social Interactionism	+	+	+	0	0	4
Napoleon complex	-	-	-	-	-	-4

“+” = positive correlation predicted; “-” = negative correlation predicted; 1 point was given for each correct prediction, 1 point subtracted for each prediction that was the opposite of the data.

#### Cognitive-Motivational-Relational

Cognitive-Motivation-Relational theory also posits “higher” cognitive functions that could be thought to act prudentially on the instigation of aggression. To the extent that this is true, stronger males would be more aggressive because of their ability endure the costs of fights. This could explain the positive correlations between strength and various measures of anger. This theory cannot explain a correlation between political aggression and strength, which could not have been formed by a prudential consideration of cost-benefit analyses. The results are partially consistent with Cognitive-Motivational-Relational theory; it scores 3 predictions out of 5 possible.

#### Cognitive Neoassociationism

Cognitive Neoassociationism predicts that negative affect creates anger before higher-order “cognitions” influence whether it is expressed or not. Given that the expression of anger has less negative consequences for stronger individuals, this theory can account for the relationship between strength and fighting history and perhaps the Utility of Personal Aggression. It cannot account for the fact that stronger males actually experience more anger (despite, presumably, experiencing less negative affect); and it cannot account for the relationship between strength and political aggression. Results are largely inconsistent with Cognitive Neoassociationism; it predicts only two of the found relationships.

#### Frustration-Aggression Model

The Frustration-Aggression Model predicts that stronger individuals should be less aggressive to the extent that they are less likely to have their goals blocked (e.g. less likely to be frustrated). This was not the case, and there is no reason to believe that stronger males were more likely to have their goals blocked; in fact the only evidence on this point – their earlier access to sexual activity and current access to a romantic partner – suggests that stronger males have less frustration. The results are inconsistent with the Frustration-Aggression Model: It could predict none of the patterns found.

#### Equity Theory

Equity theory, at least as applied to anger in general, predicts that stronger males will be less likely to feel inequitably treated and therefore less angry. The results contradict this prediction of Equity Theory as applied to human anger.

### Behaviorism

To the extent that physically stronger males were more likely to win fights and thereby be reinforced for aggression as a result, then they should be more prone to anger and aggression, at least with regard to personal conflicts of interest. Political attitudes could not have been learned in this manner, and rumination is not reinforced one way or the other. The results for personal aggression were consistent with behaviorist theory, but not for political aggression.

### Social Learning Theory

Like behaviorism, social learning theory predicts stronger males should be more aggressive as a result of their personal experience with positive reinforcement. Again, political attitudes could not have been learned in this manner, and rumination (being internal) is not subject to social learning. Results for personal aggression were consistent with social learning theory, but not for political aggression.

### Social Constructivism

If human minds are the functional outcomes of societal design one would predict that the most dangerous individuals (e.g. the strongest) would be the least likely to express aggression. The results contradict Social Constructivism.

### Social Interactionism

According to Social Interactionism, individual males will engage in conscious or unconscious cost-benefit analyses that compute the utility of various angry and aggressive responses and choose those options that give the greatest reward. This predicts that stronger males will be more prone to anger and believe that personal

aggression is an effective way of resolving disputes. The data are consistent with these predictions, except for political aggression. (Rumination is not rational, so one would expect no correlation between strength and rumination).

This theory also predicts that attitudes about political aggression should be determined, in part, by the likelihood of paying the costs of political aggression. Unlike personal aggression, political aggression (from an individual's perspective) can take place from a safe distance. Those most likely to be engaged in the conflict should, according to a rational choice model, be most against such costly action. Two items were included in the questionnaire to measure the likelihood that the subject would join the military – “I might join the armed forces” and “I think it might be very satisfying to have a career in the military.” On both measures, stronger males were more likely (though not significantly so) to indicate that they might join the military. There is no evidence that weaker males were more likely to join the military, and yet they were significantly more likely to oppose political aggression. A more direct test of this theory would be to correlate scores on the Utility of Political Aggression with the two questions regarding the likelihood of being in the military; these correlations were .064 and .156 (both  $p > .10$ , two-tailed) respectively. This contradicts the notion that aggressive attitudes are functional in that they are derived from cost-benefit calculations.

One item on the Utility of Political Aggression Scale refers to the amount of money that should be spent on the US military (“I think the US spends too much on the military,” reverse-coded). To be fair to the Social Interactionist position, this item

should be removed from the scale as a rational individual who was likely to join the military would be benefited by increasing spending on it.<sup>27</sup> The Utility of Political Aggression Scale was recoded to include only the three items that specifically and directly related to international conflict; specifically: “I think it was good for the US to show the world, by going to war against Iraq, that it has the strongest military in the world”, “I agreed more or less with the campus antiwar demonstrations”, and “I think it is a mistake to go after Middle Eastern countries that harbor terrorists. This will just lead to more attacks on America in the future.” Correlations between this scale and the subject’s likelihood of joining the military did not change for either item measuring likelihood of being in the military: .063 and .152 respectively. If anything, those who planned to join the military were more likely to support the war.

The results for political aggression cannot be explained by Social Interactionism. “Short Man Syndrome” or “Napoleon Complex”

The general prediction that physically weak males should be more angry and aggressive as a means of compensating for their strength was contradicted by the data. The more specific prediction of the model, that physically short individuals should be more prone to anger and aggression, was also empirically falsified. The correlation between height and measures of aggression is shown in Table 4-3.

---

<sup>27</sup> Though to be equally fair to critics of rational choice perspectives, bothering to compute an attitude about military spending is largely irrational as well because in a country 295 million people the probability of one man’s vote impacting military spending is preposterously low.

**Table 4-3: Correlations between Height in inches and Anger Measures**

	Proneness to Anger	Rumination	History of Fighting	Utility of Personal Aggression	Utility of Political Aggression
Height	.148 p = .252	.230 p = .072	.111 p = .388	.229 p = .074	.254 p = .046

Note: Pearson Product Coefficient, two-tailed. n = 62

As can be seen in Table 4-3, taller individuals were slightly more likely to hold attitudes about the utility of personal and political aggression. There was no evidence that shorter males were more likely to be angry as a means of overcompensation.

It should be noted that there was a .255 correlation between height and lifting strength for this sample. To determine if a general effect of physical strength was obscuring a causal relationship between height and measures of anger and aggression, a partial correlation was computed between height and measures of anger and aggression while controlling for individual's lifting strength. Results indicated no significant relationship between height and any of the measures of anger and aggression (proneness to anger, rumination, utility of personal and political aggression, and history of fighting). Thus the results are inconsistent with the "Short Man Syndrome" and the "Napoleon Complex."

## 2. Pragmatic obstacles for Study B1.

*Would these results generalize outside of a population of weight-lifters?*

It is possible that individuals who are familiar with weight-lifting equipment would be different than non-weight lifters in ways that obscure the larger relationship between strength and anger. Among such a group of people, physical strength may play a larger role in the establishment of social status, for example, and people with

higher social status should also feel entitled to a higher WTR. This possibility cannot be ruled out by this data set and will be tested in Study B2.

*Could these results have been generated by the priming effect of being in a gym during the survey portion of the experiment?*

It is possible that, because the survey was administered in the gym, that those who were stronger felt a sense of self-esteem that is not normally active when outside of the gym.<sup>28</sup> This cannot be ruled out by this data set and will be tested in Study B2.

*Possible Confound: Do angry people go to the gym more frequently?*

If angry individuals are more likely to get into physical fights, and those who engage in physical fights believe it is more important to build muscular bodies, then one would expect angry individuals to be more likely to perform weight-lifting exercises and be more likely to develop muscular bodies. Thus the relationship between strength and anger could be due solely to the fact that angry individuals build stronger bodies. While the Regulator Theory of Human Anger clearly predicts that increases in strength cause anger, it also predicts that differences in strength should lead causally to differences in anger thresholds. It is not possible to test between these ideas with the current data set, but this confound will be disentangled in Study B2.

*Are strong males more likely to turn violent but not more likely to experience anger?*

---

<sup>28</sup> Consistent with this view, Lee Kirkpatrick and colleagues (2002) has shown that individuals have different self-esteems for different domains of life and respond to attacks on those esteems differently depending on their own expertise.

The proneness to anger scale was designed to measure the experience of anger, not fighting history. Most items tap experiences, not actual fighting, but one item (If someone shoves me, I shove back) could be interpreted either metaphorically or literally, as engaging in a fight.

Nevertheless, one could argue that physically stronger males would be no more likely to experience anger, but would be more likely to use aggression to get what they want. Given that violent episodes are well-remembered, it is possible that an increased tendency to use violence (but not an increase proneness to anger) would cause an individual to perceive himself as more angry in general. This possibility cannot be ruled out entirely by this data set, but such an account does not explain the correlation between strength and political aggression. This objection will be further tested in Study B3.

***D. Study B2 - Strength predicts anger outside of the gym.***

In order to replicate and extend the findings of Study B1, a study was conducted outside of the gym with a sample of Psychology undergraduates.

The following primary hypotheses were tested again:

- 1 Stronger males should be more prone to anger.
- 2 Stronger males should NOT be more prone to angry rumination.
- 3 Stronger males should have been in more physical fights.
- 4 Stronger males should believe more in the utility of personal aggression.
- 5 Stronger males should believe more in the utility of political aggression.

**Methods for Study B2.**

Participants: Seventy-seven males (average age: 18.8) were recruited from an introductory Psychology class to participate for partial course credit.

Procedure. The procedure was the same as that used in Study B1, except there was no weight-lifting exercise (the study was conducted in the Psychology building).

Individually, each subject completed a four part procedure:

- 1). Subjects were given a questionnaire containing scales designed to measure: Proneness to Anger, Tendency to Ruminates, History of Fighting, Utility of Personal Aggression and Utility of Political Aggression. In addition, the Buss-Perry Aggression Questionnaire (Buss & Perry, 1994) was administered to subjects in an attempt to distinguish between physical aggression and anger in general. Questions regarding the frequency with which individuals attended a gym and performed strength training were also asked.
- 2). After completing the questionnaire, subjects were given a pair of black sweatpants to change into. They changed into the pants in a separate room. Then a series of morphological measurements were taken by the experimenter, including height, weight, hip circumference, waist circumference, chest circumference, unflexed bicep circumference, flexed bicep circumference, and neck circumference.
- 3). After the body measurements were taken, the subjects removed their t-shirts and were photographed from four perspectives: facing front full body, facing side full body, facing back full body, and facing forward face only.

*Strength measures*

Because this study could not be conducted at a gym, lifting strength could not be directly assessed. Instead, proxies of strength were used. There were three strength measures used in this study: flexed bicep, self-reported strength, and ratings of the individual's physical strength based on a photograph of their body. These measures were then combined into a single composite measure of strength.

### 1. Flexed Bicep Circumference

In Study B1 the best single predictor of lifting strength was flexed bicep circumference, which correlated with actual lifting strength at .735. It should be noted, however, that there is reason to believe that bicep circumference may be a better predictor of strength in gym populations than in Psychology undergraduate populations. This is because subjects in Study B1, who were recruited from the Recreation Center, appeared to have less body fat. Body fat adds to bicep circumference without increasing strength, contaminating bicep circumference as a measure of muscle width.<sup>29</sup> Thus one should not assume that the correlation between bicep circumference correlates and lifting strength is as strong in the population at large as it was in men who frequent the gym.

### 2. Self-report of physical strength

In study B1, subjects' response to the statement: "I am stronger than \_\_\_\_\_% of other men" correlated with their actual lifting strength at .643,  $p < .0001$ .

### 3. Ratings of the individuals' strength based on photographs of their bodies

---

<sup>29</sup> This is supported by the fact that subjects in Study B1 were no taller (mean height 1.81 meters for B1 and 1.79 meters for B2), not much heavier (mean weight 78.9 kilograms for B1 and 77.7 for B2), and yet had narrower hips (92.8cm for B1, 100.1cm for B2). B1 subjects also had larger average biceps (35.8 compared to 33.4).

In study B1, photographs of the subjects were rated on how strong they looked. These ratings correlated with actual lifting strength at .623,  $p < .0001$ .

In Study B2, full body photographs of the subjects wearing a standard black pair of shorts and standing next to an experimenter were taken and edited with Photoshop 8.0 so that the subject's head was covered. These full-body photographs were then shown to a sample of 50 UCSB undergraduates (18 male and 32 female, the same sample as those who rated the photographs in Study B1). The undergraduates rated the photographs on physical strength using a 7-point Likert scale rating from very weak (1) to very strong (7). Photographs were presented in random order.

#### 4. Composite measure of all proxies of strength

In study B1 the best predictor of lifting strength was a composite of all three measures listed above, each of which accounted for unique variance in lifting strength. This composite measure correlated with strength at  $r = .794$ . In Study B2, bicep circumference, self-reported strength, and the ratings of strength based on photographs of the subjects' bodies were each converted into z-scores, and these scores were averaged for a single measure of strength.

#### **Results for Study B2**

*Were the three measures of strength highly correlated?*

If these measures of strength are accurately tracking physical strength they should demonstrate high inter-correlations. Moderate correlations were found between the three strength measures as shown in Table 4-4.

**Table 4-4: Inter-correlations between measures of strength among Psychology undergraduates**

	Flexed Bicep Circumference	“I am stronger than _____% of other men.”
Flexed Bicep Circumference	1.0	.539 p < .0001
“I am stronger than % of other men.”	.539 p < .0001	1.0
Photograph Rating of Strength*	.506 p < .0001	.455, p < .0001

\*n = 77, one-tailed; for Photo Ratings of Strength n = 60 because camera malfunction prevented photographs on one day of data collection.

*Were stronger Psychology undergraduate males more prone to anger?*

Table 4-5 shows the correlation matrix with the composite measure of physical strength predicting subjects’ scores on the Proneness to Anger, Rumination, History of Fighting, Utility of Personal Aggression, and Utility of Political Aggression Scales.

**Table 4-5: Correlations between strength and attitudes on anger and aggression among Psychology undergraduates**

	Proneness to Anger	History of Fighting	Utility of Personal Aggression	Utility of Political Aggression	Rumination
Composite Measure of Strength	.219 p = .026	.330 p = .001	.233 p = .019	-.002 p = .493	.112 .163

n = 77, one-tailed (all were directional predictions except for rumination, where p values should be doubled)

Hypothesis 1: Proneness to anger

Consistent with Study B1, stronger individuals reported being more prone to anger.

#### Hypothesis 2: Rumination

Consistent with Study B1, there was no significant correlation between rumination and the composite measure of physical strength (or any single measure of strength).

#### Hypothesis 3: History of Fighting

Consistent with Study B1, scores on the History of Fighting Scale correlated significantly with physical strength.

#### Hypothesis 4: Utility of Personal Aggression

Consistent with Study B1, scores on the Utility of Personal Aggression correlated significantly with physical strength.

#### Hypothesis 5: Utility of Political Aggression

Unlike Study B1, attitudes about the efficacy of political aggression did not correlate with physical strength on the composite measure of strength (or any single measure of strength).

#### *Did stronger males score higher on the Buss-Perry Aggression Questionnaire?*

One reasonable objection to the conclusions of Study B1 is that stronger males may be more physically aggressive but no more likely to demonstrate anger outside the context of violent confrontations. The Buss-Perry Aggression Questionnaire (Buss & Perry, 1992) was derived by factor analysis to distinguish between different forms of aggression. Four factors resulted in subscales called physical aggression,

verbal aggression, anger and hostility. The physical aggression subscale contains items designed to measure the tendency to react with overt physical anger such as, “If somebody hits me, I hit back” and “Once in a while I can’t control the urge to strike another person.” The verbal aggression subscale contains items that involve one’s willingness to verbally disagree with others such as, “I tell my friends openly when I disagree with them” (relevant to the regulator theory) and “I often find myself disagreeing with people” (not relevant to the regulator theory, which makes no predictions about how often one holds non-consensual views). The anger subscale includes a hodgepodge of items that involve both one’s willingness to demonstrate annoyance such as “When frustrated, I let my irritation show” and “Some of my friends think I’m a hothead” (which are relevant to the regulator theory of anger) as well as one’s tendency to regret having gotten angry such as “Sometimes I fly off the handle for no good reason” and “I have trouble controlling my temper” (which are not relevant to the regulator theory). Finally the hostility subscale measures one’s paranoia and jealousy. Sample items include “Other people always seem to get the breaks”, “At times I feel I have gotten a raw deal out of life” and “I am suspicious of overly friendly strangers” (not relevant to the regulator theory). Only the physical aggression scale maps onto the theoretical variables of the regulator theory.

Correlations between strength measures and the four subscales of the BPAQ are shown in Table 4-6.

**Table 4-6.**

**Correlations between strength and the subscales of the BPAQ among  
undergraduate Psychology students**

	Physical Aggression	Verbal Aggression	Anger	Hostility
Composite Measure of Strength	.244 p = .015	-.077 p = .249	-.013 p = .455	-.008 p = .474

n = 77, one-tailed

Physically stronger males scored significantly higher on physical aggression.

Results on the other three subscales were inconsistent. This should not be surprising, as these scales included many questions that not relevant to the Regulator Theory.

*Is there reason to believe that stronger males are winning more conflicts of interest?*

For theoretical reasons, it is important to examine the relationship between a man's strength and his levels of frustration and likelihood of winning conflicts of interest (see discussion for Study B1). As one proxy of this relationship, questions were added to determine the sexual access available to males in the study. The following three questions were asked and correlated with the composite measure of physical strength (results are displayed in Table 4-7).

- i). "At what age did you first have sexual intercourse?"
- ii). "How many sexual partners have you had? Include sexual intercourse and oral sex."
- iii). "Are you involved in a serious romantic relationship? (yes or no)"

**Table 4-7.**

**Correlations between strength and questions about sexual access**

	Age of first sexual intercourse	Number of Sexual partners	Involvement in serious relationship
Composite Measure of Strength	-.087 p = .259	.307 p = .006	-.047 p = .342

n = 77, one-tailed; for age of first intercourse n = 58 as 19 subjects were virgins

Only the number of sexual partners was significantly correlated with physical strength. As with Study B1, there is some evidence that stronger males are more likely to win conflicts of interest than weaker males.

*Possible Confound: Do angry people go to the gym more frequently?*

A number of questions regarding gym attendance were asked in order to determine if the relationships found in Study B1 were solely due to the fact that some individuals (angrier individuals) were more likely to spend significant time in gyms and thus develop stronger bodies.

The first question designed to measure gym attendance was, “Have you ever spent more than two months working out at a gym?” Only 16 subjects out of 77 had never worked out at a gym for more than two months. An independent samples t-test was run to determine if those who had never been to a gym were less angry than those who had spent time in a gym at some point in their life. The 16 who had never worked out at a gym for at least 2 months were marginally less prone to anger ( $t(77) = 1.745, p = .08$ , two-tailed), but they showed no difference for History of Fighting (or Rumination, Utility of Personal Aggression, Utility of Political Aggression).

Two tests were run to determine if the relationships found in Study B1 were due solely to the fact that those who attend a gym are more prone to anger. The first tested the relationships in the small number of men who never attended a gym for longer than two months, and the second was a partial correlation controlling for time spent on strength training.

The data from the 16 subjects who reported never attending a gym for longer than two months were subjected to a correlational analysis between the composite measure of strength and their score on the Proneness to Anger Scale. Results showed a comparable relationship,  $r = .277$ ,  $p = .149$ . Only the last was statistically significant with  $p = .004$ . This correlation must be treated with caution given the low sample size, but demonstrates that even among those subjects who had never frequented a gym, a correlation of roughly the same effect size exists between strength and proneness to anger.

After one day of data collection it was apparent that very few subjects had never performed strength training, so an additional question was added to the survey, which asked how many months they had spent doing strength training (i.e. weight lifting). Answers ranged from 0 to 74 with a mean of 18.55 and a standard deviation of 19.52. The median time spent in strength training was 12 months. Recall that the subjects in this study were recruited from an introductory Psychology course at a liberal arts college, and yet fully half of the sample had worked at strength training for at least a year. Partial correlations were computed between the composite measure of strength and the anger-related attitudes while controlling for the number of months spent

doing strength training. Because the question measuring strength training duration was added after one day of data collection only 55 subjects were available for the analysis. The one-tailed partial correlations were as follows: Proneness to anger ( $r = .225, p = .037$ ), Rumination ( $r = .101, p = .213$ ), History of Fighting ( $r = .288, p = .010$ ), Utility of Personal Aggression ( $.222, p = .039$ ), Utility of Political Aggression ( $r = -.099, p = .218$ ). Correlations were roughly the same as those in the zero-order correlations indicating that time spent during strength training was not solely responsible for the correlation between strength and anger.

### **Discussion of Study B2**

*Did the results of Study B1 generalize outside of a gym?*

Yes. The major effects of Study B1 replicated in a more general population. The correlation between strength and the Utility of Political Aggression was not significant, but other effects did replicate significantly. One can safely rule out the possibility that strength and anger are only correlated in gym populations.

Furthermore, there was no “strength prime” in Study B2 as the experiment was conducted in a Psychology Building not a Recreation Center, and no explicit measures of strength were taken other than body measurements, which were taken after the administration of the questionnaire. One can thus rule out the possibility that the relationships in Study B1 were due to the effect of being in an environment where males compete with physical strength.

*Are strong males more likely to turn violent but not more likely to experience anger?*

Results on the Buss-Perry Aggression Questionnaire showed that physically strong males were more prone to physical aggression but not more prone to argumentation or anger (according to the Verbal and Anger Subscale of the BPAQ). This supports the view that stronger males may be more prone to aggression but not more prone to experience anger or express anger in other ways.

There is reason to doubt this interpretation, however. The anger subscale includes a number of items that refer not just to one's proneness to anger but also one's guilt over having become angry, for example, "Sometimes I fly off the handle for no good reason", and "I have trouble controlling my temper." Guilt over having become angry in a situation is not predicted by the Regulator Theory of Anger to occur more often with stronger individuals. Indeed, guilt should occur when one becomes convinced one's WTR toward a person on whom one imposed a cost was indeed too low. All else equal, the Regulator Theory predicts guilt is more likely to be correlated with being weaker, not stronger. Furthermore, items relating to anger as a personal problem in self-control are predicted to be more common among weaker individuals (note, for example, the Utility of Personal Aggression Scale, which correlated in both Study B1 and B2 with physical strength, contains items referring to anger as an effective means of conflict, not as something to be avoided). The items in the Anger Subscale suggesting that anger is a problem (as opposed to a solution) include "Sometimes I fly off the handle for no good reason", "I have trouble controlling my temper", and "I sometimes feel like a powder keg ready to explode." Being angry when you yourself feel there is no good reason, being unable to control

your reactions, and being on the verge of harming people who are not at fault (powder keg explosion) suggest dysfunction; they are not what the regulator theory means by “proneness to anger” (the regulator concept involves the probability a given cost-benefit transaction eliciting anger).

Consistent with this interpretation, the Anger Subscale of the BPAQ correlated more highly with the Rumination Scale (.448) than the Proneness to Anger Scale (.331). Furthermore, scores on the Anger Subscale of the BPAQ did not significantly correlate with the Utility of Personal Aggression (.165,  $p = .147$ ), suggesting that the Anger Subscale is not measuring an attitude that disposes the individual to feel that the violent expression of anger is effective.

The possibility that stronger males are more prone to violence but not anger has not been ruled out or confirmed by this data set. It will be tested more directly in Study B3.

*Why were the correlations between strength and attitudes about anger and aggression lower in B2 than in B1?*

The main effects of Study B1 replicated with one exception (see below), but the effect sizes were somewhat reduced. It is not clear exactly why this took place, although there are several possibilities. First, proxies of strength are less predictive of lifting strength, and they might be even less so in a population with more body fat (as in B2). Second, administering the questionnaires in the gym may have exaggerated any effect of relative strength (though it could not have created the relationship entirely). Finally subjects in B2 may have paid less attention to the questionnaire as

they were recruited for course credit not payment and the questionnaire was longer with the inclusion of the BPAQ.

The following explanations for the reduced effect sizes are not consistent with the data. 1). Truncated range in Study B2: On both strength measures and attitude measures, the Standard Deviations were comparable across both studies. 2). Lack of power in B2. There were slightly more subjects in Study B2. 3). Floor or ceiling effects; there was no indication that these were present. 4). Increased competition due to lack of romantic relationships – there is evidence that unmated males are more competitive (Daly & Wilson, 1988), but those in B1 were slightly more likely to have a serious romantic partner.

*Why did the correlation between physical strength and Utility of Political Aggression not replicate?*

No measure of physical strength significantly correlated with subjects' attitudes about political aggression in Study B2. It is unclear why this has happened. It is not due to either floor or ceiling effects in either study. The correlation between strength and attitudes about political aggression was the lowest of the effects explored in B1 and perhaps more power would reveal a more significant effect. This problem was addressed with a larger sample in Study B3.

#### **E. STUDY B3: Replication of strength and anger with a larger sample.**

In order to replicate and extend the findings of Studies B1 and B2, a study was conducted with a more diverse sample of UCSB undergraduates who were paid for their participation. Scales also had items added to them to improve reliability and

stabilize marginal effects. Finally, females were included in the study for the first time, as a control.

Seven primary hypotheses were tested:

- 1 Stronger males should be more prone to anger.
- 2 Stronger males should NOT be more prone to angry rumination.
- 3 Stronger males should have been in more physical fights.
- 4 Stronger males should believe more in the utility of personal aggression.
- 5 Stronger males should believe more in the utility of political aggression.
- 6 Stronger males should feel more entitled to better than average outcomes.
- 7 Stronger males should believe they can win more conflicts of interest.
- 8 Stronger females should not differ on anger, aggression, entitlement, or attitudes about personal or political aggression.

To address the issue of whether stronger males are not easier to anger, but simply more likely to use aggression as a convenient instrumentality, an additional scale was created to measure one's sense of entitlement to better than average outcomes. This is both of theoretical importance (as the Regulator Theory of Human Anger predicts that stronger males will seek to get better than average outcomes), but also allows one to test against the idea that stronger males may be more likely to be physically aggressive but not more prone to anger in general.

#### Individual Differences in Female Anger

The Regulator Theory of Human Anger predicts that variables that would have given females the ability to inflict costs and bestow benefits on others during human

evolutionary history would be tracked internally by females and used to regulate their Welfare Tradeoff Ratios in ways similar to physical strength in males. The question remains, however, what would have predicted a female's ability to impose costs and bestow benefits on others? It is unlikely that physical strength would have been such a predictor. Although the females of some primate species engage in physical aggression, that is not the form that competition among human females usually takes (Campbell, 2002). Historic evidence suggests that violent physical confrontation has always been the domain of males, not females (see Daly & Wilson 1988 for a review). Insofar as males interested in a female's welfare use aggression to protect and defend her, one might expect a woman's formidability to be a function of the formidability of the men she can attract. Thus it was hypothesized that a woman's physical attractiveness would predict her ability to enforce a higher WTR in others. If so, then attractiveness, and not strength, should predict proneness to anger and sense of entitlement in women.

### **Methods for Study B3**

Participants: Two hundred and eight-one subjects (125 males, average age 20; 156 females, average age 19) were recruited from the UCSB student center to participate in the study for \$10.

Procedure. Subjects completed a four part procedure:

- 1). They were given a questionnaire, which they completed in groups of 1-6. The questionnaire contained the following scales: Proneness to Anger, Tendency to Ruminant, History of Fighting, Utility of Personal Aggression, Utility of Political

Aggression, Entitlement, and the Buss-Perry Aggression Questionnaire. The scales specifically designed for this study included more items, to increase inter-item reliability and allow for more questions in case male and female reliabilities differed. Questions on the various scales were clumped, also to improve inter-item reliability. Questions regarding the frequency with which individuals attended a gym and performed strength training were also administered as well as measures of Socioeconomic Status (SES) and sexual history.

2). Subjects were taken one at a time (sometimes after their questionnaire was complete, sometimes they were asked to pause and allowed to return to the questionnaire afterwards) to a separate room where the following morphological measurements were taken by a same-sex experimenter: height, weight, hip circum., waist circum., chest circum., unflexed bicep circumference, flexed bicep circumference, shoulder-to-shoulder measure and neck circumference.

3). After the body measurements, the female subjects were given a white t-shirt to put on over their shirt and photographed from a single perspective – facing forward full-body. Male subjects were asked if they would be willing to remove their shirt for the photograph. Those willing were photographed without shirts; those who preferred to keep their shirt on were photographed in the clothing they arrived in.

#### New Items and New Scales.

New items were added to all scales except the History of Fighting Scale. In addition, two new scales were created: Entitlement and Success in Conflict. They are shown in Table 4-8.

**Table 4-8**

**New Items and New Scales for Study B3**

<p>Proneness to Anger</p>	<p>I have a short fuse          I get very angry when someone makes fun of me          If someone insults me, I usually don't say anything about it (R)          If someone gets in my face, I tell them to back off          If someone hurts my feelings I usually let it pass (R)          If someone cuts in line in front of me, I let it pass (R)          I usually shrug it off when a stranger causally insults me (R)          Sometimes I get so mad I feel like I'm going to burst          People act like jackasses all the time          People often irritate me</p>
<p>Rumination</p>	<p>It takes me a long time to get over something          I hold grudges for a long time          I sometimes wake up mad about things that happened a long time ago          Sometimes I'm seething with anger for long periods of time          After an argument I'm usually more angry than I was before          When someone makes me mad, I find myself thinking about it for weeks afterwards          When I get angry, I think it's important to act on it (R)</p>
<p>Utility of Personal Aggression</p>	<p>You have to stand up for yourself by confronting people with what they've done          You should not back down when someone threatens you; if you do back down the person will continue to take advantage of you.          Words can solve most problems better than violence. (R)          If I were to use force to solve my problems it would only cause more problems for me in the long run. (R)          If I don't fight back, people will walk all over me.          When it comes to one-on-one confrontations, violence never solves anything. (R)          Violence can solve problems for me.          Confronting people scares me. (R)          It makes me nervous to voice strong disagreement (R)</p>
<p>Utility of Political Aggression</p>	<p>In most cases I agree with the phrase, "Peace not War." (R)          When countries respond to force with force it only causes more problems in the long run. (R)          A good way for a country to protect itself is to fight harder and stronger than the opposing country.          Wars in general promote terrorism. (R)          When it comes to international conflicts, violence never solves</p>

	<p>anything. (R)</p> <p>The military can solve problems for our country.</p> <p>We need a strong military.</p> <p>To deter violence, a country needs a strong military.</p> <p>Going to war is always wrong. (R)</p>
Entitlement	<p>I deserve to have a good life.</p> <p>I don't deserve any more than anyone else. (R)</p> <p>Most people are better than me. (R)</p> <p>What I earn in life is mine, and I shouldn't be forced to share it. (R)</p> <p>I am better than most people.</p> <p>I deserve more than the average person.</p> <p>I deserve less than the average person. (R)</p> <p>People get too upset with me when I do minor things.</p> <p>When people offer to do me a favor I often refuse because I would be uncomfortable imposing on them. (R)</p> <p>I sometimes feel uncomfortable when I'm given praise. (R)</p> <p>I feel as though I need to come out on top in any confrontation.</p>
Success in Conflict	<p>If I want something, I can usually get it even if others don't want me to have it.</p> <p>Other people know not to get in my way.</p> <p>If another person and I both want something, I will be more likely to get it.</p> <p>People generally do what I ask them to do.</p> <p>I don't have much of a problem getting people to do what I want them to do.</p> <p>I can't get people to do what I want them to do. (R)</p> <p>When there's a dispute, I usually get my way.</p>

### *Strength Measures*

There were four strength measures used in this study: an inverted handgrip dynamometer, flexed bicep, self-reported strength, and ratings of the individual's physical strength based on a photograph of their body.

#### 1. Flexed Bicep Circumference

Again, flexed bicep circumference was used as a proxy for physical strength. Flexed bicep circumference correlated with lifting strength in Study B1 at .735. It is

not known the extent to which flexed bicep circumference predicts female lifting strength.

## 2. Self-report of physical strength

In study B1, subjects' response to the statement: "I am stronger than <blank>% of other men" correlated with their actual lifting strength at .643,  $p < .0001$ . It is not known how well females can estimate their strength.

## 3. Ratings of the individuals' strength based on photographs of their bodies

In study B1, photographs of the subjects were rated by Psychology undergraduates on how strong they looked. These ratings correlated with actual lifting strength at .623,  $p < .0001$ . Again, it is not known to what extent a female's physical strength can be predicted from her photograph.

In Study B3, full body photographs of the subjects standing next to an experimenter were taken and edited with Photoshop 8.0 so that the subject's head was covered. These full-body photographs were then shown to a sample of 12 UCSB undergraduates (5 male and 7 female). The undergraduates rated the photographs on physical strength using a 7-point Likert scale rating from very weak (1) to very strong (7). Photographs were presented in random order but clustered by sex (i.e. half of the subjects rated females and then males, half of the subjects rated males and then females).

## 4. "Chest Strength Dynamometer" - Inverted handgrip dynamometer.

The Rolyan Hydraulic Hand Dynamometer made by Smith & Nephew Rehabilitation is a handgrip strength tester that can measure up to 200 pounds (90-

kilograms) of force applied to its metal handle. By inverting the handgrip attachment the dynamometer can be made into a measure of chest strength.<sup>30</sup> Use of the device was demonstrated by the experimenter and then the subject demonstrated strength twice – the stronger of the two measures was taken.

##### 5. Composite Measure of Physical Strength.

In study B1 the best predictor of lifting strength was a composite of the first three measures listed above, each of which accounted for unique variance in lifting strength. That composite measure correlated with strength at  $r = .794$ . In Study B3, bicep circumference, self-reported strength, ratings of strength based on photographs of the subjects' bodies, and the Chest Strength Dynamometer were each converted into z-scores and averaged together for a single measure of strength.

##### **Results for Study B3**

*Were the measures of strength highly correlated?*

If the measures of strength are accurately tracking physical strength they should demonstrate high inter-correlations. For males, moderate correlations were found between the four strength measures as shown in Table 4-9. For females, moderate correlations exist between self-reported strength, bicep circumference, and photograph ratings of strength, but only weak relationships exist between the chest strength dynamometer measure and other measures of strength. These results suggest that the chest strength dynamometer may be a poor method of assessing female

---

<sup>30</sup> Imagine holding two bars on a prison gate and instead of trying to pull them apart, push them together. Subjects were instructed to hold the dynamometer against the chest to avoid wobbling and push as hard as they could.

strength. This could be due to the fact that, as some of the experimenters reported, many females did not appear to be trying very hard.

**Table 4-9: Inter-correlations between measures of strength among male and female undergraduates**

	Flexed Bicep Circumference	“I am stronger than ___% of other men.”	Photograph Rating of Strength
Flexed Bicep Circumference		Male: .516 Female: .416	Male: .492 Female: .514
“I am stronger than ___% of other men.”	Male: .516 Female: .416		Male: .500 Female: .417
Photograph Rating of Strength	Male: .492 Female: .514	Male: .500 Female: .417	
Chest Strength Dynamometer	Male: .552 Female: .163	Male: .536 Female: .255	Male: .462 Female: .149

n = 125 males, 156 females.

*Were stronger undergraduate males more prone to anger and fighting and more likely to believe aggression is useful?*

Table 4-8 shows the correlation matrix with the two measures of physical strength predicting subjects’ scores on the Proneness to Anger, Rumination, History of Fighting, Utility of Personal Aggression, and Utility of Political Aggression Scales.

**Table 4-8: Correlations between measures of Physical Strength and Anger among male undergraduates**

	Proneness to Anger	Rumination	History of Fighting	Utility of Personal Aggression	Utility of Political Aggression
Composite Measure of Strength	.318 p=.0001	.072 p = .194	.364 p=.00002	.323 p = .0001	.160 p = .038

n = 125, one-tailed. All were directional predictions except for rumination, where the p values should be doubled.

#### Hypothesis 1: Proneness to anger

Consistent with Studies B1 and B2, stronger individuals were more prone to anger.

#### Hypothesis 2: Rumination

There was no correlation between strength and rumination. Given no directional prediction, two-tailed tests are appropriate, so reported p values should be doubled.

#### Hypothesis 3: History of Fighting

Again, scores on the History of Fighting Scale correlated significantly with physical strength.

#### Hypothesis 4: Utility of Personal Aggression

Consistent with Studies B1 and B2, scores on the Utility of Personal Aggression correlated significantly with the composite measure of physical strength.

#### Hypothesis 5: Utility of Political Aggression

Results on the Utility of Political Aggression confirmed that stronger males were more likely to endorse political aggression as efficacious. The effect size was smaller than study B1, but was statistically significant.

*Were stronger men more likely to feel entitled and believe they could win in conflicts of interest?*

**Table 4-9: Correlations between Physical Strength and Entitlement and Success in Conflict among male undergraduates**

	Entitlement	Success in Conflict
Composite Measure of Physical Strength	.325 p = .0001	.223 p = .006

n = 125, one-tailed

Hypothesis 6: Entitlement

Stronger males were significantly more likely to feel entitled to better than average outcomes (see Table 4-9). This effect was not driven solely by the self-report measure of strength; all four components of the composite measure of strength correlated significantly with the Entitlement Scale.<sup>31</sup>

Hypothesis 7: Success in Conflicts of Interest

Stronger males were also significantly more likely to believe they can win in conflicts of interest. These two effects—positive correlations between strength and entitlement and strength and success in conflicts of interest—are important evidence for the Regulator Theory. According to the Regulator Theory, strength leads to a greater sense of entitlement (i.e., the expectation of that others will grant a higher WTR). It is this greater sense of entitlement that is hypothesized to increase the probability of anger when a given cost is imposed, and that leads to success in conflicts when higher WTRs in others make it more likely they will defer to the stronger individual.

---

<sup>31</sup> Correlations for singular measures of strength: biceps,  $r = .218$ ; self-report,  $r = .405$ ; photo ratings,  $r = .398$ ; chest strength dynamometer,  $r = .184$ .

*Did stronger males score higher on the Buss-Perry Aggression Questionnaire?*

Correlations between strength and the Buss-Perry subscales are shown in Table 4-10. As before, stronger males were more physically aggressive than weaker males. Unlike (unlike Study B2), there was a marginally significant correlation between strength and verbal aggression and a small correlation between strength and verbal aggression. The reported p values are one-tailed, but given the problems of these scales with respect to the Regulator Theory, perhaps they should be two-tailed (see discussion of Study B2 for possible problems with the Anger Subscale of the BPAQ). It is not clear whether strength *should* predict verbal aggression: if being stronger makes one feel more entitled and thus more willing to argue for one's "rights," then it would. But if being stronger causes people to defer to you without having to engage in verbal argumentation, or if weaker individuals must use verbal arguments instead of physical aggression then it should not.

**Table 4-10: Correlations between measures of strength and the subscales of the Buss-Perry Aggression Questionnaire among undergraduate UCSB students.**

	Physical Aggression	Verbal Aggression	Anger	Hostility
Composite Measure of Strength	.413 p < .000001	.177 p = .025	.142 p = .06	-.091 p = .159

n = 123, one-tailed; two subjects left several questions of BPAQ blank and were excluded from analysis.

*Is there reason to believe that stronger males are winning more conflicts of interest?*

In addition to the positive correlations found between physical strength and the Success in Conflicts of Interest scale, two questions were inserted into the survey to

measure subjects' degree of sexual access. Stronger males were less likely to be virgins ( $r = .226, p = .006$ ), had sex earlier ( $r = -.154, p = .078$ ), and had more sexual partners ( $.307, p = .0003$ ).

*Does strength predict anger in females?*

No. Because evidence suggests the psychology of women was not shaped to the same extent by an evolutionary history of frequent violent conflict, it was predicted that they would be much less likely to use the threat of physical violence as a means of increasing WTRs. Thus strength was not predicted to play as large a role in the threshold they set for anger. Because women's scores on the Chest Strength Dynamometer correlated poorly with other measures of strength, this measure was dropped from the composite measure of strength for females. Correlations between the modified Composite Measure of Strength and scores on the Proneness to Anger, Rumination, History of Fighting and Utility of Personal and Political Aggression are shown in Table 4-11.

**Table 4-11: Correlations between measures of Physical Strength and Anger among female undergraduates**

	Proneness to Anger	History of Fighting	Utility of Personal Aggression	Utility of Political Aggression	Rumination
Composite Measure of Strength	.107 p = .182	.083 p = .151	.091 p = .256	.059 p = .468	.042 p = .598

n = 156, two-tailed.

No scale of anger was significantly predicted by female strength. These null findings could not be due to a lack of power as there were 33 more women than men

in Study B3. There was no significant correlation between strength in females and any measure of anger on the BPAQ.

*Does attractiveness predict anger and sense of entitlement in females?*

Yes. Women were asked to rate how attractive they are, relative to other women at UCSB. These ratings were correlated with Proneness to Anger ( $r = .227$ ,  $p = .002$ ), Utility of Personal Aggression, ( $r = .178$ ,  $p = .014$ ), Utility of Political Aggression ( $p = .148$ ,  $p = .033$ ), Entitlement ( $r = .308$ ,  $p = .00004$ ) and Success in Conflict ( $r = .233$ ,  $p = .002$ ).

#### ***F. General Discussion for Study Set B***

According to the Regulator Theory of Anger, human males will calibrate WTRs, to some extent, based on the relative strength between themselves and others. Therefore, compared to weaker men (i) stronger men will attempt to get more resources, (ii) they will place less weight on the welfare of others in choosing their actions, and (iii) they will expect others to weigh their welfare more heavily in choosing actions that negatively impact them. Because of these differences, stronger males should be more likely to feel anger when a given cost is imposed and, all else equal, will experience more anger when a given cost is imposed. Study Set B attempted to determine if stronger males are more anger prone.

*Are stronger males more anger prone?*

Yes. Across three studies with three different subpopulations at UCSB (i.e. college gym members, undergraduate psychology majors, and undergraduates from many different majors) and across five measures of physical strength (i.e. lifting

strength, bicep circumference, self-report, ratings of the subjects' photos and chest strength dynamometer) college males showed significant correlations between strength and measures of anger, past physical violence, and attitudes about the utility of personal aggression.

*Do stronger males have a greater history of fighting?*

Importantly, strength predicted history of fighting. This should be true if the threat of aggression is a tool for negotiating WTRs, and if aggression is the functional output of the anger system when no cues of concession have been detected.

*Do stronger males feel they are more successful in negotiating situations in their favor?*

They do. The success in conflict items administered in B3 were not about physical aggression; they were more general statements about one's ability to get others to do what one wants. Stronger males felt they were better at this, which is to be expected if they expect—and are granted—a higher WTR from others.

*Do stronger males feel more entitled to resources when there is conflict?*

Yes, Study B3 showed that strength predicts a sense of entitlement. This is an important result because WTR is hypothesized to be a variable tracking how entitled an individual feels he is to deference from others in resource conflict. It is converging evidence—separate from proneness to anger or history of fighting—for the effect of strength on expected WTRs.

*Are stronger males more supportive of political aggression?*

It was predicted that stronger males would be more supportive of political aggression because the computational structure of the human mind evolved in a world with band-level conflicts in which an individual's fighting ability would have played a role in their ability to survive and benefit from warfare (Tooby & Cosmides, 1988). It is therefore reasonable, that natural selection has designed human males such that possessing greater than average physical strength leads to the adoption of attitudes about coalitional aggression that are more positive than weaker males, which are then projected on the modern world of nation states.

In Study B1 the correlation between physical strength and political aggression was significant for all measures of physical strength. This is important, because B1 had the most reliable measure of strength: actual lifting strength. This correlation did not replicate in Study B2 but did replicate in Study B3, albeit with a smaller effect size. The effects of strength were weaker for all measures in Study B2 and may have been the result of a smaller sample size and poorer scales than B3, and a weaker measure of physical strength than Study B1.

Finally, it is not surprising that the relationship between physical strength and political aggression is comparatively weak, given that decisions about collective actions must be compromises between group members. Coalitional strength (e.g. loyalty, number of warriors and their formidability, etc.) would have been a better predictor of benefiting from ancestral warfare than would personal strength – and so one would further expect the relationship between physical strength and political aggression to be smaller when coalitional strength remains uncontrolled. Consistent

with these ideas, Mackie, Devos and Smith (2000) have shown that estimates of coalition strength predict whether one is angry or fearful in cases of coalitional threat. *Are strong males more likely to use aggression instrumentally but not more likely to experience anger?*

A number of theories predict that after anger is activated a kind of cost-benefit analysis is computed designed to judge the consequences of aggressive action.<sup>32</sup> These theories could account for the relationship between physical strength and aggression as stemming from the tendency of weaker males to inhibit their aggressive responses for prudential reasons.

Study B3 suggests this is not the case for three reasons. First, the proneness to anger scale does not ask about actual aggression. Second, stronger males were not just more aggressive than weaker males, they were felt more entitled, and there is good reason to believe that a heightened sense of entitlement leads to aggressive action (see Bushman & Baumeister [2002]). Finally, in two of the three studies there was good evidence for a link between physical strength and endorsement of political aggression, which cannot be accounted for by a prudential cost-benefit analysis.

---

<sup>32</sup> Such theories include: Cognitive-Motivational-Relational, Cognitive Neoassociationism, Social Interactionism, and certain readings of Equity Theory.

## Chapter 5.

### **A New Direction in the Behavioral Output of Anger – Study C.**

The most common response to an anger-inducing event in naturally occurring situations is to engage in an argument. James Averill established this in the early 1980s with an influential study of a large sample of adults (1983). This experiment roundly contradicted the reigning theories of anger at the time, such as the Frustration-Aggression Model (Dollard, Doob, Miller, Mowrer and Sears, 1939) and Behaviorist Theory (Buss, 1961), and introduced a substantive new data point that was difficult to explain by other theories, such as Social Learning theory (Bandura, 1973). Computational models of anger at the time were too simplistic to account for complex behavioral patterns such as arguments over a provocation.

It is reasonable to conclude that the most common response to anger has been ignored because of its theoretical intractability. A theory of anger that wishes to unite causes, computations and outcomes must account for or at least provide theoretical insight into the origin, design, and function of human arguments. No theory has yet accomplished this.<sup>33</sup> This chapter explores a theoretically-driven empirical study of anger-induced arguments designed to test basic predictions from the Regulator Theory of Anger in this relatively unexplored data field.

#### ***A. The function of human arguments according to the Regulator Theory***

---

<sup>33</sup> This includes Averill's own social constructivist theory which, as was discussed in Chapter 2, provides no testable predictions about the kinds of arguments that should exist, their structure, or their function (other than a vague reference to social function).

It is theorized that human arguments serve a number of functions related to Welfare Tradeoff Ratio adjustment, and can be understood as embodying a rule-based grammar that fits those functions. The two primary functions of arguments, from the perspective of the angry individual, are as follows:

i). Recalibrate the target's Welfare Tradeoff Ratio, raising it so the target's decisions will take the angered person's welfare more into account. Successfully raising the target's WTR toward the angry individual is hypothesized to have two adaptive functions. First, it reduces the scope of cost-benefit transactions the target is willing to impose on the angry individual. Second, it raises the scope of self-beneficial cost-benefit transactions the angry individual can impose on the target without incurring anger.

The negotiative tools used include threatening to harm the target (or actually harming the target), threatening to withdraw cooperation from the target (or actually withdrawing cooperation), and (related to withdrawing cooperation) demonstrating one's value to the target (i.e., demonstrating that the target should take your welfare more into account).

ii). Change the target's estimates of the consequences, negative and positive, of their actions: leading the target to believe that the cost imposed is higher than s/he thought, and that the benefit gained was lower. The adaptive function is to prevent the target from imposing costs of that size in the future, especially for insufficient benefits.

Secondary functions most likely include:

iii). Gathering information about the magnitudes of the values involved in the cost-benefit exchange (e.g. Why did you do that?), the offender's knowledge of the magnitude of the cost, benefit and victim identity (e.g. "Do you know how much that hurt?", "Do you know who you're messing with?"), the offender's perception of their WTR with respect to you ("I thought we were friends?"), and other variables that are used to set Welfare Tradeoff Ratios.

iv). Demonstrate the lowering of the angered person's own intrinsic WTR toward the target (e.g. "Fine, if that's the way you want it, I'm not going to be your friend.")

v). Testing the boundaries of one's WTR with respect to the other and vice versa (e.g. verbal bullying).

#### 1. Recalibrate the target's Welfare Tradeoff Ratio

The primary theorized function of anger is to adjust Welfare Tradeoff Ratios in others to ensure more favorable resource divisions in the future. Much like the non-human animal competition designed for this purpose (see Chapter 1), one would predict that low cost attempts at persuasion would be done before high cost attempts, such as combat. Humans should be designed to recalibrate another's WTR without resorting to physical aggression when it would have been cost-effective to do so. It is not yet clear exactly what variables are used by individuals when setting their WTRs toward others, though there is evidence that strength plays a role in human males (see Chapter 4). Evidence from the large body of research on reciprocal altruism and mutualism also suggests that past kindness is a reason to value someone more highly (Ridley, 1998; Tooby & Cosmides, 1996; Trivers, 1971; Williams, 1966).

Thus it is predicted that among those without an inherent interest in each others' welfare (i.e. strangers, acquaintances, or those without an expectation of cooperation) arguments should contain broadcasts of strength. It is not surprising then that arguments that precede the male-male assaults and homicides that occur largely between acquaintances and strangers frequently contain threats and statements of relative fighting prowess (Felson, 1982; Luckenbill 1977), (e.g. "I can kick your ass", "come and get it, pansy", "I could beat you up one-handed").

On the other hand, when one has an expectation of inherent interest (such as with friends, kin, and others with whom long-term cooperation would be mutually beneficial) one would predict statements of relative friendship quality would be used to boost another's intrinsic Welfare Tradeoff Ratio toward oneself.<sup>34</sup> For example, "I wouldn't do that to you" or "remember when your mother was sick and I took notes for you in all your classes". Such statements can be seen as indicating the extent to which the speaker values the other person's welfare (their WTR) and is therefore useful to have as a cooperative partner. They are issued in order to demonstrate the speaker's value to the target of anger. If, for example, someone is willing to take notes for you in all your classes when your mother is sick, then they clearly value your welfare more than a stranger (who presumably would not do such a thing), and to the extent that he values your welfare, then he is in fact valuable to you and you

---

<sup>34</sup> Though not the focus of this paper, it seems likely that humans have two WTRs used toward others (Tooby & Cosmides, 2005 [HBES plenary]; see also Chapter 2). The extrinsic WTR governs cost-benefit transactions in which the individual on whom the cost was imposed is present and capable of enforcing their interests; the intrinsic WTR governs cost-benefit transactions in which the individual is not present or cannot enforce their interests. Arguments designed to change these two WTRs are hypothesized because they are set and respond to different variables.

should adjust your WTR accordingly (Tooby & Cosmides, 1996). Furthermore, these statements indicate the extent to which the target of anger would be negatively impacted by the withdrawal of cooperation. Such a threat is parallel, from a selection pressure point of view, to the threat of violence that depends on relative formidability.

These predictions are not explicitly tested in Study C, but offer a promising research vein to be explored in the future, and give theoretical insight into a host of arguments that are known to take place.

## 2. Change the target's estimates of the consequences of their actions

There are at least two reasons for angry individuals to modify the target's estimates of the cost-benefit tradeoff indicated by the target's action (e.g. how big was the cost they imposed, and how big was the benefit they received for imposing that cost). The first is to demonstrate to the target that their WTR is too low, and the other is to avoid costly actions that would result if the target of anger continued to hold misperceptions of the magnitude of the costs and benefits.

There is an inherent conflict of interest over WTRs such that others will attempt to raise yours with respect to them and resist your attempts to raise theirs with respect to you (e.g. they will want to increase the scope of cost-benefit transactions they can impose on you while resisting your cost-benefit transactions). Costly violent demonstrations or costly withdrawal of cooperation should be reserved for situations in which another individual has clearly demonstrated an unacceptable WTR. In such situations, if an angry person can demonstrate clearly that the cost-benefit tradeoff imposed on them was such that it indicated a low WTR, then there is a chance the

target of anger will spontaneously recalibrate their WTR and avoid the withdrawal of cooperation or threats of costly violence.

The second reason to recalibrate the target's perception of the magnitude of the cost they imposed on you and the magnitude of the benefit they received from imposing that cost is to avoid the future costs that would be imposed on you if such misperceptions continued. Mathematically, if someone holds a WTR toward you of  $\frac{1}{2}$ , then they would be unwilling to impose a cost of 7 on you for a benefit of 3, because the inequality for the decision rule,  $B(i) > \frac{1}{2}(C_{you})$  is not satisfied (3 is not greater than  $\frac{1}{2}(7)$ ). But if they misperceived the cost to you to actually be 5, then they would impose such a cost ( $3 > \frac{1}{2}(5)$ ). The importance of this misperception will be dependant on the likelihood of the target imposing that particular cost on you again.<sup>35</sup>

In fact, even in the absence of an explicit cost-benefit transaction, there would still have been selection to recalibrate others' perceptions of the magnitude of various costs and benefits if they were likely to result in more favorable cost-benefit transactions in the future. This offers an interesting possible explanation for one particular kind of anger that has resisted attempts at formalization – personal offense.<sup>36</sup>

---

<sup>35</sup> One might object to this line of reasoning by arguing that if someone misperceives the magnitude of a cost then one could impose that cost on those individuals for relatively small benefits and thus benefit from their misperception. This is true, but as this example demonstrates, if both individuals impose this tradeoff on each other there is a net loss of 4 to each individual (cost imposed = 7, minus benefit gained = 3). This will be the case when WTRs are 1 or less which will often be the case when it comes to intrinsic WTRs.

<sup>36</sup> Ekman and Friesen (1975) for example describes an offense to one's personal values as a source of anger; and other authors (Tedeschi & Felson, 1994; Lazarus, 1992) have referred to violations of one's "sacred values" or "social norms." It's unclear what "sacred values" refers to, but one possibility is that it refers to a disagreement over the magnitude of a cost – e.g. "was Hitler's regime really *that* bad?" If so, then offense should be a function of the likelihood of such costs being imposed on oneself

In short, the regulator theory makes precise predictions about the *content* of arguments: it predicts they should follow a particular social grammar deriving from the logic of Welfare Tradeoff Ratios. A good theory is one that can predict this content, and not just superficial characteristics such as the presence of yelling.

***B. How do other theories account for argumentation?***

1. Cognitive-Motivational-Relational

The primary motivational response of anger, according to Lazarus, is “attack.” There are appraisal components in the theory that are predicted to determine if attack is feasible or not, but at no point after the activation of anger is it predicted that individuals should attempt to modify representations in the mind of the target of anger. The existence of arguments is problematic for Cognitive-Motivational-Relational theory.

2. Cognitive Neoassociationism

It is difficult to extract testable predictions about arguments from the Cognitive Neoassociationistic model presented by Berkowitz in 1990, but the model does make explicit that “higher order cognitive processes” take place after the initial frustration or negative affect begins the activation of fear and anger. During these processes, “the affected person makes appraisals and causal attributions and considers what feelings and actions are appropriate under the particular circumstances.” During this process, presumably, the anger system could be exploited or manipulated by argumentation. Unfortunately the mechanisms for appraisal, attribution and deciding

---

and the ability to impose those costs on others (either over evolutionary time or as the result of

what is appropriate under the particular circumstances are left entirely unspecified. It is therefore impossible to derive specific predictions regarding the particulars of human arguments.

*Prediction:* Unspecified.

### 3. Equity Theory

Previous research does demonstrate that individuals are uncomfortable when they receive an inequitable share of resources (Donnerstein & Hatfield, 1982). Thus, one way of restoring equity is to demonstrate to another that their actions were inequitable and hope that, as a result of that demonstration, the individual takes steps to correct this outcome. This would predict that angry individuals should argue that the cost imposed on them was large and the benefit received by the other was large.

To the extent that the target of anger can make the angry person believe that the outcome was equitable, they can avoid the consequences of anger. Thus Equity Theory could be interpreted as predicting that the targets of anger should argue that the benefit they gained was low (e.g. "I didn't get that much") and the cost imposed on the other was low (e.g. "It didn't hurt you that bad").

*Prediction: Angry individuals should argue that the cost imposed on them and the benefit received by the other are large. Targets of anger should argue that the cost they imposed was small and that the benefit they received was small.*

### 4. Frustration-Aggression Model

---

hypotheticals run within the ontogenetic lifetime of the person).

It may be inappropriate to expect the Frustration-Aggression Model to account for human arguments as the model was designed to predict and explain instances of aggression. Such a model, however, would have to explain why most instances of the most severe cases of aggression (such as physical assault and homicide) are almost always preceded by verbal argumentation (Daly & Wilson 1988; Felson 1982; Luckenbill, 1977). In fact, the existence of verbal argumentations is counter to the logic of the Frustration-Aggression model to the extent that arguments are anything other than verbally-generated acts of aggression. If the frustration-aggression model is correct, arguments must be seen as forms of aggression and be designed to inflict harm on the other individual.

*Predictions:* Arguments should be forms of aggression the content of which is designed to cause harm.

## 5. Behaviorism

Behaviorism relies on the individual's history of reinforcement to explain everything, including verbal behavior. Such an account leaves many questions unanswered. Out of the infinite number of possible arguments, which should one make? What benefit is gained from these arguments?<sup>37</sup> How should one generalize these arguments to new situations one hasn't argued before? The unspecified nature of these problems makes it impossible to derive testable hypotheses regarding human arguments from the theory. Given that histories of reinforcement are idiosyncratic,

---

<sup>37</sup> One could argue that an apology or "winning" the argument is the unconditioned positive stimulus, but then one must specify the criteria of "winning," the nature of apologies and why some arguments elicit them more frequently than others.

behaviorism has even more problems if general patterns in the content of arguments are found. What cross-individual regularities in history of reinforcement are creating these patterns? And what is the cause of these regularities (if one rules out a domain-specialized evolved anger system that governs how people react to arguments)?

*Prediction? The existence of arguments that manifest a rule-governed cross-individual consistency is an inherent problem for Behaviorism.*

## 6. Social Learning Theory

Much like behaviorism, it is difficult to derive predictions about argumentation from social learning theory. In fact, the case for social learning theory is far worse than behaviorism in at least one respect: violence and physical aggression are almost always preceded by intense verbal arguments. Because arguments are cues that violent aggression is about to occur there would be intensely negative vicarious reinforcement for not arguing, and perhaps strong positive reinforcement for issuing sincere-sounding apologies, which almost always resolves the unpleasantness of the conflict (see Chapter 2). The existence of arguments at all is a problem for social learning theory. If one wanted to retain the theory one would be forced to argue that there are perceptible benefits that result from arguments that outweigh the negative reinforcement of violence that sometimes follow them, and that this perceptible benefit would have to follow some “kinds” of arguments more frequently than others. Finally, one would have to posit a way in which the mind categorizes arguments into abstract “kinds” because without such a categorization humans would be unable to produce sensible arguments over novel events (e.g. when a co-worker accuses you of

stepping on his foot you might reply, “but I took out the trash last night” because this argument recently won an argument with your spouse the night before).

*Prediction? The existence of rule-governed, content-structured arguments is an inherent problem for Social Learning Theory.*

#### 7. Social Constructivism

Though Averill pioneered the study of human arguments (1982, 1983) his adherence to social constructivism as an explanation for his data sets prevented him from generating testable hypotheses regarding the nature of argumentation. His data sets tended to ask “common sense” questions about the nature of arguments, rather than being theoretically driven, such as “What cues did others notice when someone was angry at them,” “how long does anger last”, “how do people feel about others after they argue with them,” or, “how do the arguments of men and women differ?” These questions, and the answers provided by his subjects, are a rich data source in need of explanation, but social constructivism is not such an explanation. As stated before, there were no predictions derived from the theory that were tested in his data set.

*Predictions: None*

#### 8. Social Interactionism

Rational Choice Theory is good at making predictions when the agents and payoffs are explicitly modeled, but in the absence of such a model it is not possible for rational choice to make good predictions about human behavior. When it comes to human arguments, it is not clear what the payoff structures are for the individuals

involved in the arguments. It is therefore not possible for rational choice (an integral part of Social Interactionism) to explain or predict what kinds of arguments should or should not exist.

I believe Tedeschi and Felson (1994) would suggest that arguments should serve the same function as aggression. There were four kinds of possible benefits from anger and aggression mentioned by Felson (2002): saving face, retaliation for past offenses, getting a resource and, possibly, entertainment. Anger-induced arguments do not appear to be done for entertainment, and even if one wanted to show that they did serve this function it is not possible to derive predictions about the kinds of arguments that should take place without a prior theory of entertainment. Arguments may be designed to save face, but again it would be impossible to derive predictions about arguments without a prior theory of face saving. It is possible that arguments are attempts to retaliate or gain a resource, but no theory is provided regarding how arguments should be structured to achieve these ends.

*Prediction: Unclear.*

### ***C. Study C: A Preliminary Analysis of Naturalistic Human Arguments***

Study C was designed to test four basic hypotheses derived from the Regulator Theory of Human Anger with regard to the basic structure of anger-induced arguments. Specifically, an angry individual on whom a cost was imposed will argue that the other individual demonstrated a low Welfare Tradeoff Ratio by arguing that 1) the cost imposed was particularly high and 2) that the benefit reaped by the target of anger was particularly low. Furthermore, the target of anger is hypothesized to resist

these accusations (so as to avoid having to change their WTR) and argue 3) that the cost imposed was particularly low and 4) that the benefit reaped particularly high.<sup>38</sup> Other predictions are also possible, but not tested below. These include arguments that the angered person is valuable to the target, that the target has been undervaluing the angered person in other cases, that the angered person may withdraw future cooperation, and that the angered person may harm the target.

### **Methods for Study C**

Participants: Thirty-nine UCSB undergraduate Psychology students (11 male and 28 female, average age: 18) were asked to recall and write down a recent argument they had engaged in, in which one or both parties were angry.

Procedure. After the thirty-nine individuals wrote about their arguments, four coders blind to the hypotheses read the arguments and attempted to determine the relevant cost-benefit transaction (i.e. what the person was angry about - what costs were imposed and benefits gained). Coders were undergraduates working for course credit and did not collaborate during the coding process. Arguments were removed from the pool if at least 3 of the 4 coders did not agree on the nature of the costs or benefits; this criterion resulted in the removal of 11 arguments.

Once the costs and benefits in the remaining 28 arguments were identified, three of the coders re-read the arguments in detail, counting each of the following types of arguments: High Cost, Low Cost, High Benefit, Low Benefit. Coders noted who made the argument and then summed the number of arguments of each type made by

---

<sup>38</sup> See Chapter 3, Experiment A3 and A4 for similar experiments with artificial arguments.

each individual. The sum of each type of argument was then averaged across raters and divided by the number of participants to yield the average number of statements used by the arguers. An example of coding derived from one subject follows.

Subject #7 engaged in an argument with her father over the following cost-benefit tradeoff: cost imposed on subject: she would not be able to work where she chooses; benefit to father: gets to keep her as a worker. She made the following high cost statement, “I would rather get paid by someone I don’t care about so I wouldn’t feel guilty for taking their money.” Her father made the following high benefit statements, “We [subject’s mother and father] need you here,” and “[You] should know it’s hard to find reliable workers.”

**Results for Study C.**

Inter-rater reliability (Cronbach’s alpha) ranged from .7 (High Benefit) to .9 (High Cost).

Table 5-1 shows the average number of statements of each type made by the angry individuals (defined as the individual on whom the cost was imposed; as it was frequently the case that both individuals expressed anger) and the target of anger.

**Table 5-1.**

**Who argued what? Average number of statements per argument made by Angry Individuals and the Targets of Anger**

	Cost to Victim		Benefit to Target	
	High	Low	High	Low
Angry Individuals	1.8	0.1	0.3	0.4

Targets of Anger	0.3	1.1	1.2	0.0
------------------	-----	-----	-----	-----

n = 28. Gray cells show statements angry individuals should make (and targets should not make); white cells show statements targets of anger should make (and angry individual should not make).

Hypothesis 1: High Cost Imposed statements should be made more often by angered individuals than by targets of anger.

A paired samples t-test was run subtracting the number of high cost arguments made by the Targets of Anger (M = 0.27, SD = .42) from the number of high cost arguments made by the Angry Individual (M = 1.79, SD = .96). The results indicated that angry individuals were significantly more likely to include High Cost statements in their arguments than the targets of anger, and the effect size was very large:  $r = .83$  ( $t(27) = 7.64, p < .0001, \text{one-tailed}$ )

Hypothesis 2: Low Benefit Gained statements should be made more often by angered individuals than by targets of anger.

A paired samples t-test was run subtracting the number of low benefit arguments made by the Targets of Anger (M = .03, SD = .14) from the number of low benefit arguments made by the Angry Individual (M = 0.36, SD = .87). Results indicated that, as predicted, angry individuals were significantly more likely to include Low Benefit statements in their arguments than the targets of anger ( $r = .36; t(27) = 2.00, p = .027, \text{one-tailed}$ ).

Hypothesis 3: High Benefit Gained statements should be made more often by targets of anger than by angry individuals.

A paired samples t-test was run subtracting the number of high benefit arguments made by the angry individual ( $M = 0.32$ ,  $SD = .59$ ) from the number of high benefit arguments made by the target of anger ( $M = 1.17$ ,  $SD = 1.23$ ). Results indicated that, as predicted, targets of anger were significantly more likely to include High Benefit statements in their arguments than the angry individuals,  $r = .55$  ( $t(27) = 3.38$ ,  $p = .001$ , one-tailed).

Hypothesis 4: Low Cost Imposed statements should be made more often by targets of anger than by angry individuals.

A paired samples t-test was run subtracting the number of low cost arguments made by angry individuals ( $M = 0.12$ ,  $SD = .23$ ) from the number of low cost arguments made by the target of anger ( $M = 1.13$ ,  $SD = .93$ ). Results indicated that, as predicted, targets of anger were significantly more likely to include Low Cost statements in their arguments than the angry individuals, with large effect size  $r = .72$  ( $t(27) = 5.40$ ,  $p < .0001$ , one-tailed).

### **Discussion of Study C**

Strong support was found for all four hypotheses. Angry individuals made arguments designed to indicate that the other individuals demonstrated a low Welfare Tradeoff Ratio (i.e. High Cost Imposed and Low Benefit Gained statements). The targets of anger resisted these statements and made arguments indicating that their actions did not indicate a low Welfare Tradeoff Ratio (i.e. Low Cost Imposed and High Benefit Gained statements).

These results cannot be explained as attempts by the angry individual to demonstrate inequity, because angry individuals tended to argue that the other individual received a low benefit (thus increasing the *equity* of outcomes, compared to a situation in which the target had received a high benefit). They are not consistent with the target of anger attempting to demonstrate equity either, because the targets of anger tended to argue that they received a high benefit (thus arguing for greater inequity of outcome than if they had received a lower benefit).

The results could not have been predicted by Behaviorism, Social Learning Theory, Cognitive Neoassociationism, Social Interactionism, Frustration-Aggression Model, or a social constructivist position.

The importance of human arguments in anger research has been vastly underestimated. Rather than treating arguments as societal constructions that were designed to circumvent the “innate” aggressive response of anger (Averill, 1982; Izard, 1977)<sup>39</sup>, one might look at arguments as focused manipulations of key variables that are used by the anger mechanism – such as calculations of the magnitude of costs and benefits. Much of the data gathered by Averill’s extensive survey of instances of anger is consistent with this interpretation. For example, on page 202, Averill (1982) shows that anger resulted in reinterpretations of relevant variables for 62% of the instances of anger he analyzed. Most commonly, the angry person reinterpreted the motives (i.e. decided the target received a high benefit) or the guilt of the target of anger. The second most common reinterpretation by the angry individual was that he

or she “decided the incident was less important than originally thought” (i.e. low cost to themselves). The target of anger also responded in ways consistent with the Regulator Theory. According to Averill (1982, pg. 223), the most common long-term response on behalf of the target of anger was that he or she, “realized [his/her]] own faults.” Fully 76% of the eighty subjects reported that they “somewhat” or “very much” agreed with that statement. But only 39% of subjects, in contrast, “did something that was good for the angry person” in response to the argument. This difference highlights that anger tends to calibrate another’s mental representations rather than evoke restitution.

Restitution does play a role in the Regulator Theory of Anger: it is one way the target can honestly signal having raised his/her WTR toward the angered individual. Consistent with this, using a cooperative game drawn from experimental economics, Bottom et al. (2002) have show that apologies are more likely to elicit future cooperation when they are accompanied by some form of restitution.

According to the regulator theory, after an argument, targets can signal a modified WTR toward the angered individual in one of two ways.

(1) Impose a cost on oneself to benefit the angered individual (e.g. give them something you value). The higher the cost incurred and the lower the benefit provided, the greater the signaled WTR. For example, by walking 5 miles just to bring Jan the bag lunch she forgot, Fred has signaled a very high WTR toward Jan—higher than that signaled by Bill, who walked 5 miles to bring Alice the powerpoint file she

---

<sup>39</sup> It should be noted that Averill would distinguish his view from Izard’s. Averill holds that the whole

needed for an important job interview. The powerpoint Bill provided may be a larger benefit to Alice, but it does not signal a larger WTR than Fred's bringing the lunch.

(2) Refuse to take a benefit that would impose a cost on the angered individual. The higher the benefit forgone and lower the cost that would have been imposed, the higher the signaled WTR. Thus refusing to accept a free vacation because it would mean the new guy at work would have to handle the customers himself signals a higher WTR than refusing to accept the free vacation because it would lead to the new guy being deported for not having a job. Note that refusing to accept a cost-imposing benefit is not "recompensing" the angered individual at all. That is, one need not provide a compensating benefit to signal an improved WTR. One can signal an improved WTR by refusing to impose a cost.

---

of anger is socially constructed – not just the argumentative responses.

## Chapter 6.

### Conclusions and unanswered questions.

This final chapter reviews the primary data points established in the empirical studies presented herein and evaluates the ability of alternative theories of anger to account for their existence. Finally, an attempt is made to explain the failure of anger theories to account for the rich datasets that anger theorists have discovered, and an appeal made to use evolutionary theory to derive more comprehensive and theoretically-unified theories to account for additional, currently undiscovered, features of anger.

#### *A. New Features of Anger*

Seven different theories were described in Chapter 2 as alternative to the Regulator Theory of Anger. Each of the theories had difficulty accounting for what I listed as the six major features of anger. Those that could account for them most clearly (Cognitive-Motivational-Relational theory and Social Interactionism) had unspecified intuitive concepts that drove much of the theoretical model, including its predictions (e.g. “intentionality,” “demeaning offense,” “face”, “rational,” etc.).

To that list of features I would add the replicated and reliable effects of Study Set A, Study Set B and Study C.

1. Feature #7: Anger becomes increasingly more likely to result from a cost-benefit transaction as the benefit to target of anger decreases. (See Studies A1 and A4).

2. Feature #8: Individual differences in the proneness to anger, in males but not females, are correlated with physical strength

It should be noted that this effect is specific to anger proneness, and not to rumination, hostility, or ability to control unreasonable anger. Moreover, strength also predicts actual fighting history.

The Regulator Theory of Anger predicts that increases in strength will lead causally to increases in anger, but this feature has not been empirically verified. Furthermore, there is now some evidence that this attitude also appears for political conflicts, something that no other theory of anger predicts. Because this evidence is weaker than the other forms of aggression it is not included as a stable, reliable feature of anger – perhaps future research will be more conclusive.

3. Feature #9: There appears to be a grammar to human arguments, consistent with the hypothesis that arguments are designed to manipulate perceptions of Welfare Tradeoff Ratios. See Studies A3 and C.

Angered individuals argue that the cost-benefit transaction imposed on them indicated a low WTR on the other's part (i.e. that the other individual did not sufficiently value the welfare of the angered individual) and the targets of anger argued that the WTR expressed by the cost-benefit transaction was higher (i.e. that they did sufficiently value the welfare of the angry individual).

### ***B. How do other Theories account for New Features of Anger***

The ability of other theories of anger to predict and explain these effects varies and is detailed, as best I can, in Table 6-1. Recall the coding scheme used: +2 – the

theory clearly predicts the feature and the absence of the feature would be evidence against the theory; +1 – the theory is consistent with the data point but does not require it; 0 – the theory makes no prediction about the feature; -1 – the theory is contradicted by the evidence for that feature of anger.

**Table 6-1: Predictions of current models against new data**

	Effect of Benefit	Strength and Anger	Arguments	Total score
Regulator Theory	2	2	2	6
Cognitive Motivational Relational	1	1	-1	1
Cognitive Neoassociationism	0	-1	0	-1
Equity Theory	-1	-1	-1	-3
Frustration Aggression	1	-1	-1	-1
Behaviorism	1	1	-1	1
Social Learning	1	1	-1	1
Social Constructivism	0	-1	0	-1
Social Interactionism	1	2	0	3

1. Cognitive-Motivational-Relational Theory

Lazarus’s model of anger relies on a number of intuitive concepts that could, if explicitly detailed, account for the effect of the offender’s benefit on the cause of

anger. He refers to “blameworthiness” as an appraisal that must be computed by the human mind early in the anger process (1992, pg. 223). Without a computational model of blameworthiness however, it remains a post-hoc explanation that cannot predict which actions will be blameworthy and which will not. For example, is imposing a cost more blameworthy when it was done for large gain or for small gain? If the purpose of blame is to motivate actions that would deter the particular offense, imposing costs for large gains would need to be more blameworthy. Indeed, this seems to be implicit in the criminal justice system, where stealing a lot leads to a longer sentence than stealing a little (all else equal). Lazarus also suggests that anger is more likely when the offense is, “best ameliorated by attack.” Given that stronger males face less negative consequences from ‘attacking’, the Cognitive-Motivational-Relational theory of anger could be seen to predict that stronger males will aggress more frequently and have a lower threshold for anger.<sup>40</sup> Lazarus has a very hard time, however, accounting for the existence and details of arguments. Lazarus claims, “few would argue with the proposition that the innate action tendency in anger is *attack* on the agent held to be blameworthy for the offense,” (pg. 226). He also claims that this action tendency is inhibited for “personal or social reasons.” What these unspecified reasons replace “attack” with is not clear, nor is why a system designed to attack is so easily overridden or satisfied with argumentation that has little to do with attacking.

---

<sup>40</sup> While it would be hard to argue that stronger males faced more consequences during an evolutionary history without a police state; it is not clear to me that stronger males in current environments are rewarded for their aggressive behavior. They are, of course, more likely to “win” the fight, but may be more likely to be punished for blamed for the behavior. At the least, this data point should be explored by those who adhere to a more rational-choice explanation of differential anger in stronger and weaker males, as it is crucial to their account.

It would be as if an organism that was deprived of food would simply breathe harder to compensate.

## 2. Cognitive Neoassociationism

The logic of Cognitive Neoassociationism (Berkowitz, 1990) does not make explicit specific predictions about any of the effects of anger. It does predict that anger-related feelings, action tendencies, thoughts and memories are triggered by a variety of circumstances, but which thoughts, actions, memories, and feelings are not specified (other than the fact that they are related to anger, of course). As such, the role of the offender's benefit is not deemed explicitly relevant. Individual differences in anger, according to this view, should stem largely from 1) individual differences in the tendency to experience negative affect, and 2) prudential factors (e.g. the anticipation of punishment). Stronger males presumably experience less negative affect (see Study B1-B3), but they would presumably anticipate fewer negative consequences from expressing anger. According to Cognitive Neoassociationism, the contemplation of anticipated consequences of aggressive behavior occurs *after* anger is triggered. Thus stronger males should experience less anger, but may react with physical aggression more frequently given the activation of anger. Finally, this perspective makes no predictions about the nature of arguments.

## 3. Equity Theory

Equity theory, as applied to anger (Donnerstein & Hatfield, 1982)<sup>41</sup> predicts that as the inequity between two individuals grows they should be more likely to become angry. As the benefit to the offending individual increases so does the inequity, and thus, the anger. This is not how anger responds to cost-benefit transactions (see Studies A1, A4 and C). Equity theory is less specific about the role of stronger males; but to the extent that stronger males are receiving disproportionately high shares of valuable resources they should be prone to guilt, not anger. This is also not true (see Studies B1-B3). Finally, human arguments could be interpreted by Equity Theory to be attempts to signal inequity in hopes of arousing a sense of guilt that would lead to recompense. If this were the case, then individuals who were angry should argue that the cost imposed on them was large and the benefit to the other was large. This is not the case (see Study C).

#### 4. Frustration-Aggression Model

Later versions of the Frustration-Aggression Model posited that “arbitrary” frustrations were more likely to elicit aggression than “non-arbitrary” ones. These were variously characterized as “unjustified”, “controllable” or “unexpected.” In most of the examples labeled arbitrary, the value of the contested resource to the offender has been raised. Thus, one could subsume the effect demonstrated in Study Set A (that anger is less when the offender received a larger benefit for imposing the cost) under the umbrella of “arbitrary” – though with a loss of computational specificity.

---

<sup>41</sup> I should note once again that Donnerstein & Hatfield recognized some of the shortcomings of this approach, and they should be given great credit for proposing an idea with such specific predictions, even given its problems. The traditional solution to specific but incorrect theories appears to be vague

The fact that stronger males are more prone to anger is difficult for the Frustration-Aggression Model to account for because stronger males are less likely to be frustrated (e.g. they have more sexual partners and feel they can more easily win conflicts of interests). Finally, the Frustration-Aggression Model posits that aversive stimuli result in aggression. The existence of arguments, let alone their conformity to a model of animal conflict, is an inherent problem for the Frustration-Aggression Model.

#### 5. Behaviorism and Social Learning Theory

Buss (1961) and Bandura (1973) applied behaviorist theory to the study of aggression. These theories make no specific predictions about the role of the offender's benefit and appear to be pliable in either direction. One could imagine that those who benefit the most from an offense will be the hardest to deter (e.g. if someone hurts you for 7 units of hedonic value, you would have to punish them  $>7$  units to deter that action, thus protecting yourself). This could mean that those who gain more should be punished more (and thus cause more anger) or that those who gain less should be punished more because there is a better chance of deterring them. Behaviorism and Social Learning Theory are somewhat more specific when it comes to the correlation between strength and anger. Stronger males would, presumably, be less likely to experience negative outcomes for the expression of anger, and thus

---

but unfalsifiable theories, and Donnerstein & Hatfield should be given credit for avoiding that approach.

behaviorism can account for the fact that stronger males are more aggressive.<sup>42</sup>

Social Learning Theory could maintain this explanation, though to the extent that aggressive behavior is modeled after others, the person's individual physical strength becomes much less important – and thus social learning theory has a harder time accounting for this finding.<sup>43</sup> Finally, both Behaviorism and Social Learning theory are ill-equipped to explain the existence or form of human arguments.

#### 6. Social Constructivism

Averill (1982, 1983) advocated a social constructivist explanation for anger and aggression. Despite his thorough review of historical perspectives on anger, and his data rich research efforts, his theory cannot make specific predictions about the computational structures that identify cues that trigger anger, such as the role of the offender's benefit. One could tentatively predict from a social functional perspective, that stronger males should be less prone to anger and aggression (given the harm that is visited upon society by such individuals). This is not true (see Study Set B). Finally, despite his pioneering effort into the realm of human arguments, social constructivism can not make predictions about the existence of arguments, their structure, the degree of their cross-cultural generality, their variance between friends and strangers, or any other feature.

#### 7. Social Interactionism

---

<sup>42</sup> Again, this assumes that stronger males are less likely to be punished for their aggression (or that the punishment is outweighed by the benefits of winning the fight). This is possible, but not known to be true.

<sup>43</sup> It could be rescued by positing that stronger males are more likely to model their behavior after other stronger males who (because of operant conditioning) are more aggressive.

Tedeschi and Felson make no explicit reference to the relative valuation of a resource in their theory of coercive action. But they do refer to three functions of such action, “(a) to influence others to obtain some benefit, (b) to express grievances and establish justice, and (c) to assert or defend social identities.” Without a computational definition of “justice” or “social identities” it is difficult to derive predictions from the second two functions. The first could make different predictions about the role of the benefit – 1) the more another values a contested resource the more anger is required to get them to give you the resource (thus it should increase anger), but 2) the less another values a contested resource the more likely anger is to get them to give you the resource (thus it should decrease anger). It is therefore hard for the theory to make a prior prediction, but they can account for the effect retrospectively. Social Interactionism, like many theories of anger, posits that individuals make sensitive cost-benefit calculations about the use of aggression. As such, they can account for the effect of stronger individuals increase in aggression and anger – “people are more likely to use aggression against those who they believe have less coercive power than themselves,” (Felson, 2002, pg 54). Finally, Social Interactionism does not specify (and it is difficult to extract) predictions regarding the structure and nature of human arguments (see Study C discussion section).

***C. The Regulator Theory accounts for the known features of anger***

1. Feature #1: Anger frequently results from the imposition of larger costs

One’s Welfare Tradeoff Ratio with respect to another individual regulates the kinds of cost-benefit transactions one is willing to impose on them and what cost-

benefit transactions one is willing to tolerate from them. Low Welfare Tradeoff Ratios (e.g. valuing that individual's interests less) lead one to impose larger costs on that individual for smaller benefits. Holding the benefit constant, the larger the cost an individual knowingly imposes on you, the lower their Welfare Tradeoff Ratio and the more likely anger is to be triggered.

### 2. Feature #2: Intentionally increases the likelihood of anger

A Welfare Tradeoff Ratio is consulted when individuals make planned actions that affect the welfare of others. If the offender did not make such a plan, the costs-benefit transaction that results would not indicate anything about the individual's Welfare Tradeoff Ratio toward you. Beyond explaining the effect of intentionality, the Regulatory Theory offers the beginnings of a computationally specific theory of intentionality. Namely that intentionality should contain at least these three components: i) knowledge of the magnitude of the cost, ii) knowledge of the magnitude of the benefit, and iii) knowledge of the target's identity (see Experiment A2).<sup>44</sup>

### 3. Feature #3: Apologies mitigate anger

Just as natural selection has designed animals to respond to "submission" displays by ceasing conflicts of assessment (i.e. fights over Resource Holding Power), it

---

<sup>44</sup> Because one's actions often affect the welfare of many different people (e.g. deciding to move to a new city has impact on friends, family, co-workers, children, children's friends, children's friends' parents, store owners, neighbors, etc.), it seems likely that humans first consider those for whom they have a higher WTR first (e.g. spouse, children), and ignore the interests of those for whom one has a low intrinsic WTR (e.g. the paperboy who has one less customer). It is therefore predicted that failure to consider the impact of one's actions on another does indicate something about one's WTR toward them; especially if one did consider the welfare of another - such neglect is predicted to cause anger

should also design the anger mechanism to respond to signals of recalibration of both Welfare Tradeoff Ratios and perceptions of resource value. Apologies appear to serve that function.

Not only does the Regulator Theory account for the effect of apologies, but it introduces a new way of exploring the nature of apologies. At a minimum, there ought to be two kinds of apologies: recalibrations of Welfare Tradeoff Ratios (e.g. “I’m sorry, you deserve better than that”), and recalibrations of the magnitude of cost-benefit transaction (e.g. “I’m sorry, I had no idea that would hurt you so much.”).<sup>45</sup> The Regulator Theory also makes predictions about the kinds of actions that qualify as demonstrating a “real” apology, specifically actions that demonstrate a high Welfare Tradeoff Ratio. For example, forgoing a high benefit to avoid imposing a small cost on the person is an indicator of a high WTR, as is imposing a large cost on yourself in order to give them a small benefit.

#### 4. Feature #4: Anger and aggression are often used by males to restore “face”

A large subset of physical fights among human males can be understood as contests over relative formidability, just as conspecific combat in other animals is frequently designed to estimate relative RHP (see Chapter 1). Human fights parallel these animal contests in many ways including: clear signals of aggressive intent

---

(e.g. an angry spouse yells, “you asked your boss about moving but you didn’t bother to ask me what I wanted?!”)

<sup>45</sup> An interesting case of anger is predicted to arise when someone demands a WTR apology when you feel a resource-valuation apology is sufficient. For example, you take your friend’s last soda, thinking he has more. He is thirsty and demands that you apologize; you say you’re sorry and point out that you didn’t know it was his last. He says he believes that you didn’t know, but continues to say, “that’s not good enough.” His continued anger even when you’ve recalibrated the cost is a signal that you are not allowed to take a soda even if it had not been his last. Such denial of a benefit to a friend can be

(often in the form of insults, see below), body posturing that exaggerates body size, escalations of aggression that start with low cost demonstrations of strength (e.g. shoving, pushing, wrestling) that escalate to more violent demonstrations if neither individual submits (e.g. biting, gouging, weapon use), clear signals of submission, and individual differences in propensity to aggress that track life history variables in functional ways (e.g. young, unmated males show highest propensity for aggression).

5. Feature #5: Personal insults are one of the most reliable causes of anger

Just as other animals have clear signals of aggressive intent, so are humans predicted to have low-cost means of i) provoking a conflict of assessment (e.g. a fight over who is “tougher”) and ii) indicating disagreement about a resource division.

Insults appear to serve the first function, and perhaps the second as well.

Insults are further predicted to set the parameters of the disagreement so that the target of the insult knows what is being disputed. For example, imagine a man who is working late into the night and calls his wife to get her advice about a problem. She is sleeping when he calls and is upset that he has woken her. If she were to call him, “careless” or even “stupid,” she would be implying that he did not appreciate the cost to her of waking her up. If she called him a “bastard,” or an “asshole,” she would be implying that he did not value her enough and was demonstrating a low Welfare Tradeoff Ratio. Insults are predicted to distinguish between extrinsic and intrinsic WTRs are well. For example, if an acquaintance says you’re being a “dick” or a “bastard” because you borrowed his laptop without asking, he is saying something

---

anger provoking, and the Regulator Theory would predict that you would be angry in this situation

about your intrinsic WTR. If he calls you a “pussy” (man) or a “whore” (woman) he is saying something about your extrinsic WTR – and would presumably be signaling an intent or willingness to enforce his interests with social or physical violence.

6. Feature #6: Anger has cross-culturally universal features and neuro-physiological locality

Anger is an adaptation designed by natural selection and that its basic computational structure should be universal across cultures. Many of the universal features of anger serve the same function as these features do in other animals: signals of resource contention and submission (e.g. facial and vocal expressions), and physiological preparedness for aggression if necessary. The cross-cultural similarities in the antecedents of anger are also predicted by the Regulator Theory.

The Regulator Theory predicts many other universal features of anger that have not yet been documented, including all the major effects of Study Sets A and B, and Study C.<sup>46</sup>

For a more thorough review of the Regulator Theory’s account of these six features of anger, see Chapter 2.

#### ***D. Conclusions: How to account for complex order***

Anger is complex.

Social science has had difficulty accounting for complex features of human behavior largely because it has ignored the one known cause of complex functional

---

even though no cost was imposed on you, in fact quite the opposite.

<sup>46</sup> This is not to say that there should be no cultural differences in anger. See Tooby & Cosmides (1992) for an evolutionary psychological approach to understanding culture.

design in organisms, natural selection (Tooby & Cosmides, 1992). Without functional theories capable of generating testable predictions about many different aspects of the domain of study, researchers have had two choices: 1) retreat into smaller data sets that can be predicted and cogently summarized by one or two main effects, or 2) posit intuitive concepts to account for a great range of data but remain computationally intractable and pliable enough to account for shifting datasets.

It is otherwise difficult to account for why five extremely intelligent scientists from five diverse fields of study attempted to account for all aggressive behavior in terms of unexpected goal blockage in 1939 (Dollard, Doob, Miller, Mowrer & Sears).<sup>47</sup> Certainly clinical psychologist Dr. Dollard must have noticed that some people use aggression to get rewards that they didn't previously expect (e.g. crimes of opportunity); social psychologist Dr. Doob must have noticed that insults and inappropriate requests are a large cause of aggression; physiologist Dr. Miller must have noticed that facial and vocal expressions and physiological changes that occur along with aggression were clearly designed by some system that would play a role in the behavioral aspects of aggression; learning theorist Dr. Mowrer, being of a behaviorist frame of mind, must have noticed that aggressive behavior could be "conditioned" to respond to stimuli other than frustration; developmental theorist Dr. Sears surely noticed that children get some spontaneous joy out of aggressive play, or use aggression to bully others. A comprehensive theory to account for all these behaviors was missing – a restricted dataset, however, could be explained.

A similar restriction of the dataset is required to explain why Berkowitz (1990) declared that, “negative affect of any kind will first activate anger-related feelings, action tendencies, and thoughts and memories.” Surely Berkowitz was aware that depression held a weak relationship to anger at best – in fact in his 1990 paper he attempts to defend the relationship between depression and anger with references to the fact that infants often conflate the anger and sad face, those who mourn the death of a loved one will occasionally display anger, and psychoanalysts view depression as being caused by anger. The defense seems odd given the massive difference in the tendency for anger and sadness to trigger aggression.<sup>48</sup> The theory cannot account for that broad gap with ease and turns to the second way to avoid positing complex functional systems—appeal to vague constructs that can allow one to channel one’s intuition and retrospectively account for a vast array of datasets – “this conception assumes that associative networks link specific types of feelings with particular thoughts and memories and also with certain kinds of expressive-motor and physiological reactions.”

Theories designed to account for broad ranges of data on anger and aggression have a difficult time uniting all the different features known to exist into a simple principle without positing abstract concepts that are intractable. Social constructivism, for example, resists the ability to make predictions but can readily account for almost any data set that is found. Cognitive-Motivational-Relational

---

<sup>47</sup> Dollard was a clinical psychologist, Doob a social psychologist, Miller a physiological psychologist, Mowrer a learning theorist, and Sears a developmental psychologist.

<sup>48</sup> Lazarus (1992), for example, argues quite reasonably that the primary action pattern for sadness is “inaction.”

theory and Social Interactionism both point (rightly, in my opinion) to mental states of others as the primary cause of anger, but they lack a theoretical background capable of making specific predictions. Hence the authors of the theories make sweeping reviews of the theoretical literature and use intuitive concepts such as “justice,” “demeaning offense,” “social identity,” etc., as accounts for those effects.

The alternative to vague theories, or theories that are specific but restricted to smaller data sets, are computationally-specific, functional theories that posit many testable hypotheses based on a simple model of selection pressures and the logical extensions of them given what is known of human evolutionary history. Natural selection is the only process shown capable of organizing and designing organisms. It is the only process that could have designed anger,<sup>49</sup> and thus any functional design in the system (including the reliably developing facial expressions, vocalizations, physiological changes, recalibrational learning mechanisms, and behavioral responses) can only be summarized by computational designs that could have solved the selection pressures that created anger.

It is a daunting task to map a human emotion. The known features of anger are staggeringly large,<sup>50</sup> and the features that are unknown are almost certainly more numerous. Anger is most likely designed to work differently when one is angry at strangers, friends, enemies, kin, children, men and women. The effects anger has on the perceptual systems are unknown but likely substantial. The nature of insults is

---

<sup>49</sup> Or, if one prefers to theorize that anger is the result of learning, natural selection is the only process that could have designed the learning mechanism responsible for learning anger.

<sup>50</sup> I have listed 9 of them here but each can be broken into many parts and there are no doubt many others.

almost completely unexplored. There are a number of causes of anger that, while nominally referenced, have no current explanation: violations of “sacred” values, another person’s anger directed at you (e.g. “how dare you get angry at me!”, Ekman & Friesen, 1975), or inappropriate requests. The role of vocal changes in anger, although detailed, are not explained. The role of recurrent offenses (e.g. “the first time was okay, but now it’s making me angry”) is unexplored. Intentionality remains to be fully explored and mapped computationally such that it can be modeled without intuitive references. What it means to be a hypocrite, and why it is so anger provoking, is unanswered.

Anger is complex. But complex functional systems can be understood when their function is identified. There are precedents – the mapping of the visual system (Marr, 1982), human universal grammar (Pinker, 1994), mechanisms of social exchange (Cosmides & Tooby 1989), and, of course, the evolution of fighting systems in non-human animals (Archer, 1988; Huntingford & Turner, 1987). The selection pressures described in this paper may or may not be the ones most responsible for the functional structure of human anger, though they have had great early success in making predictions and explaining many of the most reliable and significant features of anger. Ultimately, the Regulator Theory of Anger will be judged by its ability to generate computationally-specific explanations capable of organizing and explaining features of anger that have been intractable by the intuitive nature of previous theories and the restricted datasets that characterize much of the human anger research. If it fails, the alternative must be another computationally-specific, functional theory that

is consistent with evolutionary biology. The dataset on anger is now too large to be described by simple learning mechanisms and too detailed to be accounted for by vague theories that do not computationally specify their primary components.

## References

- Adelmann, P. & Zajonc, R. (1989). Facial efference and the experience of emotion. *Annual Review of Psychology*, 40, 249-280.
- Alexander, R. (1961). Aggressiveness, territoriality, and sexual behavior in field crickets (*Orthoptera: Gryllidae*). *Behaviour*, 17, 130-223.
- Archer J. (1988). *The behavioural biology of aggression*. Cambridge: Cambridge University Press.
- Archer, J., Birring, S. & Frederick, W. (1998). The association between testosterone and aggression among young men: empirical findings and a meta-analysis. *Aggressive Behavior*, 24, 411-420.
- Arnold, Magda. 1960. *Emotion and personality*. New York: Columbia University Press.
- Austad, S. (1983). A game theoretical interpretation of male combat in the bowl and doily spider. *Animal Behaviour*, 31, 59-73.
- Averill, J. (1982). *Anger and aggression: An essay on emotion*. New York: Springer-Verlag.
- Averill, J. (1983). Studies on anger and aggression: Implications for theories of emotion. *American Psychologist*, 38, 1145-1160.
- Axelrod, R. & Hamilton, W. (1981). The evolution of cooperation. *Science*, 211, 1390-1396.
- Balikci, A. (1970). *The Netsilik Eskimo*. Garden City, NY: The Natural History Press.
- Bandura, A. (1973). *Aggression: A social learning analysis*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1983). Psychological mechanisms of aggression. In R. G. Geen & E. I. Donnerstein (Eds.), *Aggression: Theoretical and empirical reviews* (Vol. 1, pp. 1-40). San Diego, CA: Academic Press.

- Barki, A., Harpaz, S. & Karplus, I. (1997). Contradictory asymmetries in body and weapon size, and assessment in fighting male prawns, *Macrobrachium rosenbergii*. *Aggressive Behavior*, 23, 81-91.
- Baron, R. & Bell, P. (1977). Sexual arousal and aggression by males: effects of type of erotic stimuli and prior provocation. *Journal of Personality and Social Psychology*, 35, 79-87.
- Banse, R. & Scherer, K. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70(3), 614-636.
- Berg, I. & Fox, V. (1947). Factors in homicides committed by 200 males. *Journal of Social Psychology*, 26, 109-119.
- Berkowitz, L. (1962). *Aggression: a social psychological analysis*. New York: McGraw-Hill.
- Berkowitz, L. (1968). Impulse, aggression and the gun. *Psychology Today*. Vol. 2(4), 19-22.
- Berkowitz, L. (1989). Frustration-aggression hypothesis: Examination and reformulation. *Psychological Bulletin*, 106, 59-73.
- Berkowitz, L. (1990). On the formation and regulation of anger and aggression: a cognitive-neoassociationistic analysis. *American Psychologist*, 45(4), 494-503.
- Berkowitz, L. & Walster, E. (1976). (Eds.) *Advances in experimental social psychology* (vol. 9). New York: Academic Press.
- Blumstein, A. & Cohen, J. (1980). Sentencing of convicted offenders: an analysis of the public's view." *Law and Society Review*, 14, 223-261.
- Boehm, C. (1984). *Blood revenge. The anthropology of feuding in Montenegro and other tribal societies*. Lawrence, KS: University Press of Kansas.
- Borden, R., Bowen, R. & Taylor, S. (1971). Shock-setting behavior as a function of physical attack and extrinsic reward. *Perceptual and Motor Skills*, 33(2), 563-568.
- Bottom, W., Gibson, K., Daniels, S. & Murnighan, K. (2000). When talk is not cheap: Substantive penance and expressions of intent in rebuilding cooperation. *Organization Science*, 13, 497-513.
- Brace, R. & Pavey, J. (1978). Size-dependent dominance hierarchy in the anemone *Actinia equina*. *Nature*, 273, 752-753.

- Bramel, D. (1969). Interpersonal attraction, hostility, and perception. In J. Mills (Ed.), *Experimental social psychology* (pp. 1-120). New York: Macmillan.
- Brown, D. (1991). *Human Universals*. New York: McGraw-Hill.
- Bruce, R. H. (1941). An experimental analysis of social factors affecting the performance of white rats. III. Dominance and cooperation motivated by water and food deprivation. *Journal of Comparative Psychology*, *31*, 395.
- Bushman, B & Baumeister, R. (2002). Does self-love or self-hate lead to violence?. *Journal of Research in Personality*, *36*(6), 543-545.
- Buss, A. (1961). *The psychology of aggression*. New York: Wiley.
- Buss, A. (1966). Instrumentality of aggression, feedback, and frustration as determinants of physical aggression. *Journal of Personality and Social Psychology*, *3*, 153-162.
- Buss, A. & Perry, M. (1994). The aggression questionnaire. *Journal of Personality and Social Psychology*, *63* (3), 452-459.
- Campbell, A. (2002). *A mind of her own: The evolutionary psychology of women*. London: Oxford University Press.
- Cannon, W. (1929). *Bodily change in pain, hunger, fear and rage: An account of recent researches into the function of emotional excitement*. New York: Appleton-Century-Crofts.
- Chagnon, N. (1983). *Yanomamo: The Fierce People*. (3<sup>rd</sup> ed.). NY: Holt, Rinehart and Winston.
- Chloe, J. C. (1994). Sexual selection and mating system in *Zorotypus gurneyi* Choe (Insecta: Zoraptera). *Behavioral Ecology and Sociobiology*, *34*, 87-93.
- Clutton-Brock, T. & Albon, S. (1979). The roaring of red deer and the evolution of honest advertisement. *Behaviour*, *69*, 145-170.
- Cohn, A. (1955). Social norms, arbitrariness of frustration, and status of the agent of frustration in the frustration-aggression hypothesis. *Journal of Abnormal and Social Psychology*, *51*, 222-226.
- Cole, D. D. & Shaffer, J. N. (1966). A study of social dominance in cats. *Behaviour*, *27*, 39.

- Connell, J. H. (1963). Territorial behavior and dispersion in some marine invertebrates. *Researches on Population Ecology*, 5, 87-101.
- Cortes, J. B. & Gatti, F. M. (1972). *Delinquency and crime: a biopsychological approach*. New York: Seminar Press.
- Cosmides, L. (1989). The logic of social exchange: has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, 31, 187-276.
- Cosmides, L. & Tooby, J. (1987). From evolution to behavior: Evolutionary psychology as the missing link. In J. Dupre (Ed.), *The latest on the best: Essays on evolution and optimality*. Cambridge, MA: MIT Press.
- Cosmides, L. & Tooby (1989). Evolutionary psychology and the generation of culture, Part II. A computational theory of social exchange. *Ethology and Sociobiology*, 10, 51-97.
- Cosmides, L. & Tooby, J. (1994). Beyond intuition and instinct blindness: toward an evolutionary rigorous cognitive science. *Cognition*, 50(1-3), 41-77.
- Cosmides, L. & Tooby, J. (2000). Consider the source: The evolution of adaptations for decoupling and metarepresentations. In D. Sperber (Ed.), *Metarepresentations: A multidisciplinary perspective* (pp. 53-115). Vancouver Studies in Cognitive Science. NY: Oxford University Press.
- Craig, K., O'Neal, E., Taylor, L., Yost, E., Langley, T., Rambow, R., Allgoewer, A. & Folger, R. (1993). Equity and derogation of those against whom we have aggressed. *Aggressive Behavior*, 19, 355-36.
- Cristol, D. (1992). Food deprivation influences dominance status in dark-eyed juncos, *Junco hyemalis*. *Animal Behaviour*, 43, 117-124.
- Dabbs, J. & Dabbs, M. (2000). *Heroes, rogues, and lovers: testosterone and behavior*. New York, NY, US: McGraw-Hill.
- Daly, M. & Wilson, M. (1983). *Sex, evolution, and behavior* (2<sup>nd</sup> edition). Belmont: Wadsworth Publishing Company.
- Daly, M. & Wilson, M. (1988). *Homicide*. New York: de Gruyter.
- Davies, N. & Halliday, T. (1978). Deep croaks and fighting assessment in toads *Bufo bufo*. *Nature*, 274, 683-685.

- Dengerink, H. & Bertilson, H. (1974). The reduction of attack instigated aggression. *Journal of research in personality*, 8(3), 254-262.
- Dixon, K. & Cade, W. (1986). Some factors influencing male-male aggression in the field cricket *Gryllus integer* (time of day, age, weight and sexual maturity). *Animal Behaviour*, 34, 340-346.
- Dollard, J., Doob, N., Miller, N., Mowrer, O. & Sears, R. (1939). *Frustration and aggression*. New Haven, CT: Yale University Press.
- Donnerstein, E., & Barrett, G. (1978). The effects of erotic stimuli on male aggression towards females. *Journal of Personality and Social Psychology*, 36, 180-188.
- Donnerstein, E. & Hatfield, E. (1982). Aggression and Inequity. In Greenberg, J. & Cohen, R. (Eds.) *Equity and justice in social behavior*. New York: Academic Press.
- Dowds, B. & Elwood, R. (1985). Shell wars: assessment strategies and the timing of decisions in hermit crab shell fights. *Behaviour*, 85, 1-24.
- Dunlop, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, 1, 170-177.
- Eberhard, W. (1979). The functions of horns in *Podischnus agenor* (Dynastinae) and other beetles. In M. S. Blum & N.A. Blum (Eds.), *Sexual Selection and Reproductive competition in Insects*. Pp. 231-258. Academic Press, London.
- Ekman, P. (1973). Cross-cultural studies of facial expression. In P. Ekman (Ed.), *Darwin and facial expression: A century of research in review* (pp. 169-222). New York: Academic Press.
- Ekman, P. (1994). Antecedent events and emotion metaphors. In Ekman, P. & Davidson, R. (Eds.) *The nature of emotion: fundamental questions*. Oxford: Oxford University Press.
- Ekman, P. & Friesen, W. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17(2), 124-129.
- Ekman, P., Friesen, W. (1975). *Unmasking the face*. Englewood Cliffs, NJ: Prentice-Hall.
- Ekman, P., Levenson, R. & Friesen, W. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, 221(4616), 1208-1210.

- Ekman, P. & Oster, H. (1979). Facial expressions of emotion. *Annual Review of Psychology*, 30, 527-554.
- Enquist, M. & Leimar, O. (1983). Evolution of fighting behaviour; decision rules and assessment of relative strength. *Journal of Theoretical Biology*, 102, 387-410.
- Enquist, M., Leimar, O., Ljungberg, T., Mallner, Y. & Segerdahl, N. (1990). A test of the sequential assessment game: fighting in the cichlid fish *Nannacara anomala*. *Animal Behaviour*, 40, 1-14.
- Epps, P., & Parnell, R. W. (1952). Physique and temperament of women delinquents compared with women undergraduates. *British Journal of Medical Psychology*, 25, 249-255.
- Epstein, S. (1965). Authoritarianism, displaced aggression and social status of the target. *Journal of Personality and Social Psychology*, 2, 585-589.
- Epstein, S. & Taylor, S. (1967). Instigation to aggression as a function of degree of defeat and perceived aggressive intent of the opponent. *Journal of personality*, 35(2), 265-289.
- Evans, M. & Hatchwell, B. (1992). An experimental study of male adornment in the scarlet-tufted malachite sunbird: I. The role of pectoral tufts in territorial defence. *Behavioural Ecology and Sociobiology*, 29, 413-419.
- Ewing, L. S. (1967). Fighting and death from stress in a cockroach. *Science*, 155, 1035-1036.
- Faruzzi, A., Solomon, M., Demas, G. & Huhman, K. (2005). Gonadal hormones modulate the display of submissive behavior in socially defeated female Syrian hamsters. *Hormones and Behavior*, 47(5), 569-575.
- Fehr, E. and 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.
- Felson, R. (2002). *Violence and gender reexamined*. Washington DC: American Psychological Association.

- Felson, R., Baccaglioni, W., & Gmelch, G. (1986). Bar-room brawls: Aggression and violence in Irish and American bars. In A. Campbell & J.J.Gibbs (Eds.), *Violent transactions* (pp. 153-166). Oxford, England: Basil Blackwell.
- Fessler, D. & Haley, K. (2003). The strategy of affect: emotions in human cooperation. In Hammerstein, P. (Ed.) *Genetic and Cultural Evolution of Cooperation* (pg. 7-36). Cambridge: MIT Press.
- Frantz, C. & Benningson, C. (2005). Better late than early: The influence of timing on apology effectiveness. *Journal of Experimental Social Psychology*, 41(2), 201-207.
- Freedman, J. (1975). *Crowding and behavior*. San Francisco: Freeman.
- Freud, S. (1963). Instincts and their vicissitudes. In P Rieff (Ed.), *Freud: General psychological theory* (pp. 83-103). New York: Collier.
- Frijda, N. (1994). Lex Talionis: On Vengeance. In Goozen, S., Van de Poll, N., & Sergeant, J. (Eds.) *Emotions: Essays on Emotion Theory*. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Galati, D., Sini, B., Schmidt, S. & Tinti, C. (2003). Spontaneous facial expressions in congenitally blind and sighted children aged 8-11. *Journal of Visual Impairment & Blindness*, 97(7), 418-428.
- Gallistel, R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.
- Geen, R. (1968). Effects of frustration, attack and prior training in aggressiveness upon aggressive behavior. *Journal of Personality and Social Psychology*, 9(4), 316-321.
- Geen, R. (1998). Aggression and antisocial behavior. In Gilbert, D. & Fiske, S. (Eds). *The handbook of social psychology, Vol. 2 (4th ed.)*. New York, NY, US: McGraw-Hill.
- Geen, R. & Quanty, M. (1977). The catharsis of aggression: An evaluation of a hypothesis. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol .10, pp. 2-39). San Diego, CA: Academic Press.
- Gibbens, T. C. N. (1963). *Psychiatric Studies of Borstal Lads*. London: Oxford University Press.
- Glueck, S., and Glueck, T. T. (1956). *Physique and Delinquency*. New York: Harper.

- Ginsberg, B. & Allee, W. (1942). Some effects of conditioning on social dominance and subordination in inbred strains of mice. *Physiological Zoology*, 15, 485-506.
- Goffman, E. (1971). *Relations in public*. New York: Basic Books.
- Greenwell, J. & Dengerink, H. (1973). The role of perceived versus actual attack in human physical aggression. *Journal of Personality and Social Psychology*, 26(1), 66-71.
- Guhl, A. M. (1956). The social order of chickens. *Science in America*, 194(2), 42-46.
- Hale, E. B., (1956). Effects of forebrain lesion on the aggressive behavior of Green Sunfish, *Lepomis cyanellus*. *Physiological Zoology*, 29(2), 107 – 127.
- Hamilton, V. L., and Rytina, S. (1980). Social consensus on norms of justice: should the punishment fit the crime? *American Journal of Sociology*, 85, 1117-1144.
- Hammerstein, P. & Parker, G. A. (1982). The asymmetric war of attrition. *Journal of Theoretical Biology*, 96, 647-682.
- Hammock, G., Rosen, S., Richardson, D., & Bernstein, S. (1989). Aggression as equity restoration. *Journal of Research in Personality*, 23(4), 398-409.
- Hansen, A. J. (1986). Fighting behavior in bald eagles: a test of game theory. *Ecology*, 67, 787.
- Harrell, W. (1980). Retaliatory aggression by high and low Machiavellians against remorseful and non-remorseful wrongdoers. *Social Behavior and Personality*, 8, 217-220.7
- Harrell, W. & Hartnagel, T. (1979). Aggression against a remorseful wrongdoer: The effects of self-blame and concern for the victim. *Journal of Social Psychology*, 107, 267-275.
- Harris, MB. (1993). How provoking: What makes men and women angry? *Aggressive behavior*, 19(3), 199-211.
- Hazlett, B. A. (1966). Factors affecting the aggressive behaviour of the hermit crab *Calcinus tibicen*. *Zeitschrift fuer Tierpsychologie*, 23, 655.
- Hazlett, B. A., Rubenstein, D. I. & Ritschoff, D. (1975). Starvation, aggression, and energy reserves in the crayfish *Orconectes virilis*. *Crustaceana* 8, 11.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

- Helm, B., Bonoma, T. & Tedeschi, J. (1972). Reciprocity for harm done. *Journal of Social Psychology*, 87, 89-98.
- Henderson, D. & Chiszar, D. (1977). Analysis of aggressive behaviour in the bluegill sunfish *Lepomis macrochirus rafinesque*: effects of sex and size. *Animal Behaviour*, 25, 122-130.
- Herrnstein, R. (1977). The evolution of behaviorism. *American Psychologist*, 32, 593-603.
- Hofmann, H. A. & Schildberger, K. (2001). Assessment of strength and willingness to fight during aggressive encounters in crickets. *Journal of Animal Behaviour*, 62, 337-348.
- Holtgraves, T. (1989). The form and function of remedial moves: reported use, psychological reality, and perceived effectiveness. *Journal of Language and Social Psychology*, 8, 1-16.
- Huntingford, F. A., and Turner, A. K. (1987). *Animal conflict*. Chapman & Hall, New York.
- Izard, C. (1977). *Human emotions*. New York: Plenum Press.
- Jacoby, S. (1983). *Wild Justice*. New York: Harper & Row.
- Jakobsson, S., Radesaeter, T. & Jaervi, T. (1979). On the fighting behaviour of *nannacara anomala* (Pisces, Cichlidae) males. *Zeitschrift for Tierpsychologie*, 49, 210-220.
- Jaffe, Y., Malamuth, N., Feingold, J. & Feshbach, S. (1974). Sexual arousal and behavioral aggression. *Journal of Personality and Social Psychology*, 30(6), 759-764.
- Jegard, S. & Walters, R. (1960). A study of some determinants of aggression in young children. *Child Development*, 31, 739-747.
- Jensen, P. & Yngvesson, J. (1998). Aggression between unacquainted pigs-- sequential assessment and effects of familiarity and weight. *Applied Animal Behaviour Science*, 58(1), 49-61.
- Johnstone, T. & Scherer, K. (2000). Vocal communication of emotion. In Lewis, M. & Haviland-Jones, J. (Eds.) *Handbook of emotions 2<sup>nd</sup> edition* (pp. 220-235).

- Kane, T., Joseph, J. & Tedeschi, J. (1977). Perceived freedom, aggression, and responsibility, and the assignment of punishment. *The Journal of Social Psychology*, 103, 257-263.
- Karsten, R. (1935). *The head-hunters of western Amazonus*. Helsingfors: Societas Scientiarum Fennica.
- Katano, O. (1990). Dynamic relationships between the dominance of male dark chub, *Zacco temmincki*, and their acquisition of females. *Animal Behaviour*, 40(6), 1018-1034.
- Kaufmann, H. (1970). *Aggression and altruism*. New York: Holt, Rinehart & Winston.
- Keeley, L. (1996). *War before civilization*. Oxford: Oxford University Press.
- Kelly, R. (1995). *The foraging spectrum*. Washington and London: Smithsonian Institution Press.
- King, H. (1961). Psychological effects of excitation in the limbic system. In D. E. Sheer (Ed.), *Electrical stimulation of the brain*. Austin: University of Texas Press.
- Krebs, J. & Davies, N. (1993). *An introduction to behavioural ecology (3rd edition)*. Oxford: Blackwell Scientific Publications.
- Lack, D. (1966). *Population studies of birds*. Oxford: Clarendon Press.
- Lazarus, R. (1991). *Emotion and adaptation*. Oxford: Oxford University Press.
- Lazarus, R. (1994). Appraisal: the long and the short of it. In Ekman, P. & Davidson, R. (Eds.) *The nature of emotion: fundamental questions* (pp. 208-215). Oxford: Oxford University Press.
- LeBlanc, S. (2003). *Constant battles: the myth of the peaceful, noble savage*. New York: St. Martin's Press.
- LeBoeuf, B. (1972). Sexual behaviour in the northern elephant seal *Mirounga angustirostris*. *Behaviour*, 41, 1-26.
- LeBoeuf, B. (1974). Male-male competition and reproductive success in elephant seals. *American Zoologist*, 14, 163-176.
- Lewis, O. (1961). *The children of Sanchez: Autobiography of a Mexican family*. NY: Random House.

- Lorenz, K. (1966). *On aggression*. New York: Harcourt, Brace and World.
- Luckenbill, D. (1977). Criminal homicide as a situated transaction. *Social Problems*, 25, 176-186.
- Kirkpatrick, L., Waugh, C., Valencia, A. & Webster, G. (2002). The functional domain specificity of self-esteem and the differential prediction of aggression. *Journal of Personality and Social Psychology*, 82(5), 756-767.
- Mackie, D., Devos, T. & Smith, E. (2000). Intergroup emotions: Explaining offensive action tendencies in an intergroup context. *Journal of Personality and Social Psychology*, 79(4), 602-616.
- Maier, R. (1964). The role of the dominance-submission ritual in social recognition of hens. *Animal Behaviour*, 12, 59.
- Maier, R. (1998). *Comparative Animal Behavior: an Evolutionary and Ecological Approach*. Boston: Allyn and Bacon.
- Marques, D. & Valenstein, E. (1977). Individual differences in aggressiveness of female hamsters: response to intact and castrated males and to females. *Animal Behaviour*, 25, 131-139.
- Marr, D. (1982). *Vision: a computational investigation into the human representation and processing of visual information*. W.H. Freeman and Co.
- Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.
- Maynard Smith, J. & Parker, G. A. (1976). The logic of asymmetric contests. *Animal Behavior*, 24, 159-175.
- Mitchell, K. (1976). Competitive fighting for shells in the hermit crab, *Clibanarius vittatus*. *Aggressive Behavior*, 2, 31-37.
- Neil, S. (1985). Size assessment and cues: studies of hermit crab contests. *Behaviour*, 92, 22-38.
- Nickel, T. (1974). The attribution of intention as a critical factor in the relation between frustration and aggression. *Journal of Personality*, 42, 482-492.
- Nisbett, R. & Cohen, D. (1996). *Culture of honor: The psychology of violence in the South*. Boulder, CO: Westview.

- Nowlis, V. (1941). The relation of degree of hunger to competitive interaction in chimpanzee. *Journal of Comparative Psychology*, 32, 91.
- Ohbuchi, K. & Kambara, T. (1985). Attacker's intent and awareness of outcome, impression management, and retaliation. *Journal of Experimental Social Psychology*, 21, 321-330.
- Ohbuchi, K., Kambara, T. & Agarie, N. (1989). Apology as aggression control: Its role in mediating appraisal of and response to harm. *Journal of Personality and Social Psychology*, 56, 219-227.
- O'Leary, M. & Dengerink, H. (1973). Aggression as a function of the intensity and pattern of attack. *Journal of Research in Personality*, 7(1), 61-70.
- Owens, D. & Owens, M. (1996). Social dominance and reproductive patterns in brown hyaenas, *Hyaena brunnea*, of the central Kalahari desert. *Animal Behaviour*, 51(3), 535-551.
- Panksepp, J. (2000). Emotions as natural kinds within the mammalian brain. In Lewis, M. & Haviland-Jones, J. (Eds.) *Handbook of emotions 2<sup>nd</sup> edition* (pp. 137-156).
- Parker, G. A. (1974). Assessment Strategy and the Evolution of Fighting Behaviour. *Journal of Theoretical Biology*, 47, 223-243.
- Parker, G. (1978). Searching for mates. In J. R. Krebs & N.B. Davies (Eds.), *Behavioural Ecology: An Evolutionary Approach*, 1<sup>st</sup> edition, pp.214-244. Blackwell Scientific Publications, Oxford.
- Pastore, N. (1952). The role of arbitrariness in the frustration-aggression hypothesis. *Journal of Abnormal and Social Psychology*, 47, 728-731.
- Petrie, M. (1984). Territory size in the moorhen (*Gallinula chloropus*): an outcome of RHP asymmetry between neighbours. *Animal Behaviour*, 32, 861-870.
- Pinker, S. (1994). *The Language Instinct: How the Mind Creates Language*. New York: William Morrow and Company.
- Poole, J. (1999). Signals and assessment in African elephants: evidence from playback experiments. *Animal Behaviour*, 58, 185-193.
- Preston, K. & Kimberley S. (1987). "What's the worst thing...?" Gender-directed insults. *Sex Roles*, 17(3), 209-219.

- Ridley, M. (1998). *The origins of virtue: human instincts and the evolution of cooperation*. London: Penguin Books.
- Rime, B. & Giovannini, D. (1986). The physiological patterns of reported emotional states. In Scherer, K., Wallbott, H. & Summerfield, A. (Eds.) *Experiencing Emotion: A cross-cultural study* (pp. 84-97). Cambridge: Cambridge University Press.
- Riordan, C., Marlin, N. & Kellogg, R. (1983). The effectiveness of accounts following transgression. *Social Psychology Quarterly*, *46*, 213-219.
- Robitaille, J. & Bovet, J. (1976). Field observations on the social behaviour of the Norway rat, *Rattus norvegicus* (Berkendhout). *Biology of Behaviour*, *1*, 289-308.
- Rohwer, S. & Rohwer, F. (1978). Status signaling in Harris sparrows: experimental deceptions achieved. *Animal Behaviour*, *26*, 1012-1022.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.
- Sargent, W. (1974). *People of the valley*. NY: Random House.
- Schlenker, B. (1980). *Impression management: The self-concept, social identity and interpersonal relations*. Monterey, CA: Brooks/Cole.
- Schwartz, G., Kane, T., Joseph, J., & Tedeschi, J. (1978). The effects of post-transgression remorse on perceived aggression, attribution of intent, and level of punishment. *British Journal of Social and Clinical Psychology*, *17*, 293-297.
- Scott, J. & Fredericson, E. (1951). The causes of fighting in mice and rats. *Physiological Zoology*, *24*, 273-309.
- Shasby, G., and Kingsley, R. F. (1978). A study of behavior and body type in troubled youth. *Journal of School Health*, *48*, 103-107.
- Shaver, P., Schwartz, J., Kinon, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology*, *52*, 1061-1086.
- Sheldon, W. (1940). *Varieties of Human Physique*. New York: Harper & Brothers Publishers.
- Sheldon, W. H., Hartl, E. M. & McDermott, E. (1949). *Varieties of Delinquent Youth*. New York: Harper.

- Sheldon, W. H., Dupertius, C. W. & McDermott, E. (1954). *Atlas of Men: A Guide for Somatotyping the Adult Male at All Ages*. New York: Harper.
- Shortell, J., Epstein, S. & Taylor, S. (1970). Instigation to aggression as a function of degree of defeat and the capacity for massive retaliation. *Journal of Personality*, 38(3), 313-328.
- Shostak, M. (1981). *Nisa: the life and words of a !Kung woman*. Cambridge: Harvard University Press.
- Siegel, A. & Schubert, K. (1995). Neurotransmitters regulating feline aggressive behavior. *Reviews of Neuroscience*, 6, 47-61.
- Sinclair, A. (1977). *The African Buffalo*. University of Chicago Press, Chicago.
- Smuts, B., Cheney, D. Seyfarth, R. Wrangham, R. & Struhsaker, T. (1987). *Primate societies*. Chicago: The University of Chicago Press.
- Stenberg, C. R.; Campos, J. J.; Emde, R. N. (1983). The facial expression of anger in seven-month-old infants. *Child Development*, 54(1), 178-184.
- Stenberg, C. R., & Campos, J. J. (1990). The development of anger expression in infancy. In N. Stein, B. Leventhal, & T. Trabasso (Eds.), *Psychological and biological approaches to emotion* (pp. 247-282). Hillsdale, NJ: Lawrence Erlbaum.
- Taylor, S. (1967). Aggressive behavior and physiological arousal as a function of provocation and tendency to inhibit aggression. *Journal of Personality*, 35(2), 297-310.
- Tedeschi, J. & Felson, R. (1994). *Violence, aggression & coercive actions*. Washington DC: American Psychological Association.
- Toch, H. (1969). *Violent men*. Chicago: Aldine.
- Tokarz, R. (1985). Body size as a factor determining dominance in staged encounters between male brown anoles (*Anolis sagrei*). *Animal Behaviour*, 33, 746-753.
- Tooby, J. (1982). Pathogens, polymorphism and the evolution of sex. *Journal of Theoretical Biology*, 97, 557-576.
- Tooby, J. & Cosmides, L. (1988). The Evolution of War and its Cognitive Foundations. *Institute for Evolutionary Studies Technical Report 88-1*.

- Tooby, J. & Cosmides, L. (1992). The psychological foundations of culture. In J. Barkow, L. Cosmides, & J. Tooby (Eds.) *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.
- 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. & DeVore, I. (1987). The reconstruction of hominid behavioral evolution through strategic modeling. In Kinzey, W. (Ed.) *The evolution of human behavior: Primate models*. SUNY Press, Albany.
- Trivers, R. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46, 35-57.
- Tyler, S. J. (1972) The behaviour and social organization of the New Forest Ponies. *Animal Behaviour Monographs*, 5, 85-196.
- Wallbot, H., Ricci-Bitti, P. & Baenninger-Huber, E. (1986). Non-verbal reactions to emotional experiences. In Scherer, K., Wallbott, H. & Summerfield, A. (Eds.) *Experiencing Emotion: A cross-cultural study* (pp. 98-116). Cambridge: Cambridge University Press.
- Wallbot, H., Scherer, K. (1986). The antecedents of emotional experiences. In Scherer, K., Wallbott, H. & Summerfield, A. (Eds.) *Experiencing Emotion: A cross-cultural study* (pp. 69-83). Cambridge: Cambridge University Press.
- Walster, E. Berscheid, E., & Walster, G. (1973). New directions in equity research. *Journal of Personality and Social Psychology*, 25, 151-176.
- Warr, M., Gibbs J. P., and Erickson, M. L. (1982). Contending theories of criminal law: statutory penalties versus public preferences. *Journal of Research in Crime and Delinquency*, 19, 25-46.
- Weiss, W. (1969). Effects of the mass media of communication. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (Vol. 5, 2<sup>nd</sup> ed., pp. 77-195). Reading, MA: Addison-Wesley.
- 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, G. (1966). *Adaptation and Natural Selection*. Princeton, NJ: Princeton University Press.

Williams, G. (1975). *Sex and Evolution*. Princeton, New Jersey: Princeton University Press.

Williams, J. (1980). *Duelling in the old south. Vignettes of social history*. College Station, TX: Texas A & M University Press.

Wolfgang, M. E. (1958). *Patterns in criminal homicide*. Philadelphia, PA: University of Pennsylvania Press.

Worchel, S., Arnold, S. & Harrison, W. (1978). Aggression and power restoration: The effects of identifiability and timing on aggressive behavior. *Journal of Experimental Social Psychology*, 14, 43-52.

Yarrow, L. (1948). The effects of antecedent frustration on projective play. *Psychological Monographs*, 62,(6, Whole No. 293).

Zack, S. (1975). A description and analysis of agonistic behavior patterns in an opisthobranch mollusk, *Hermisenda crassicornis*. *Behaviour*, 53, 238-267.

Zechmeister, J., Garcia, S., Romero, C. & Vas, S. (2004). Don't Apologize Unless You Mean It: A Laboratory Investigation of Forgiveness And Retaliation. *Journal of Social and Clinical Psychology*, 23(4), 532-564.

Zillman, D. (1971). Excitation transfer in communication-mediated aggressive behavior. *Journal of Experimental Social Psychology*, 7, 419-434.

## Appendix A: Scenarios for Experiments A3 and A4

Experiment A3: Scenarios and statements were counterbalanced across subjects.

Imagine you went to Mr. Peterson's pharmacy and shoplifted some prescription drugs. A week later Mr. Peterson is watching surveillance videos, and sees you shoplifting. When you come back to the store, he grabs you and threatens to call the police. He is very angry with you. You admit that you took the drugs, but you want to argue that what you did was not that bad. How much would each argument strengthen or weaken your case? (Rate each statement as if it is the only argument you are going to make on your own behalf.) Assume that he will believe you when you make the statement.

- \_\_\_\_\_ You used the medicine to treat a potentially lethal infection
- \_\_\_\_\_ You had stolen before
- \_\_\_\_\_ You used the medicine to treat a mild acne problem
- \_\_\_\_\_ You had enough money in your wallet to pay for the medicine
- \_\_\_\_\_ The theft had cost Mr. Peterson a great deal of money
- \_\_\_\_\_ Two other people had robbed the store earlier that year
- \_\_\_\_\_ The theft had cost Mr. Peterson very little money

Imagine you and a woman named Sandra work at a local restaurant by the beach. Sandra is mad at you for talking on the phone when you were supposed to be bussing tables. She had to work your tables as well. She is mad at you, but you don't think you did anything very bad. How much would each argument strengthen or weaken your case? Assume that Sandra will believe you when you make the statement.

- \_\_\_\_\_ You were talking to your girlfriend and finally convinced her to marry you during that phone call
- \_\_\_\_\_ Sandra is twenty-two years old
- \_\_\_\_\_ You were talking to someone you were dating casually, but didn't like much
- \_\_\_\_\_ You didn't realize how long you were talking.
- \_\_\_\_\_ There were lots of customers that day
- \_\_\_\_\_ You knew how long you were talking.
- \_\_\_\_\_ There were not many customers that day

Imagine that you ruined one of your friend's sweaters. It is completely useless to him now, and he is very upset about it. You want to argue that what you did was not that bad. How much would each argument strengthen or weaken your case? Assume that he will believe you when you make the statement.

- \_\_\_\_\_ You ruined his sweater on purpose.
- \_\_\_\_\_ You ruined his sweater by using it to stop the bleeding from a gash on your head caused by a car crash.
- \_\_\_\_\_ You didn't ruin his sweater on purpose.
- \_\_\_\_\_ You ruined his sweater by using it to clean your tub.
- \_\_\_\_\_ You didn't know it was his sweater when you ruined it.
- \_\_\_\_\_ His grandmother had knit him the sweater before she died.
- \_\_\_\_\_ You knew it was his sweater when you ruined it.
- \_\_\_\_\_ He got the sweater at a garage sale for a dollar.

Scenarios for Experiment A4: Scenarios and statements were counterbalanced across subjects.

Imagine you own a pharmacy, and someone shoplifted some prescription drugs. You discover this while watching surveillance videos a week after the theft. One day you see the thief come into the store, and you grab him and threaten to call the police. He admits to the theft, and you are very angry with him, but he doesn't think what he did was that bad. How much would each fact affect your anger? (Rate each fact as if it is the only one from the list that you are considering.)

- \_\_\_\_\_ The guy used the medicine to treat a potentially lethal infection
- \_\_\_\_\_ The guy had enough money in his wallet to pay for the medicine
- \_\_\_\_\_ The guy used the medicine to treat a mild acne problem
- \_\_\_\_\_ Two other people had robbed the store earlier that year
- \_\_\_\_\_ The theft had cost you a great deal of money
- \_\_\_\_\_ The guy had stolen before
- \_\_\_\_\_ The theft had cost you very little money

Imagine you and a woman named Sandra work at a local restaurant by the beach. You are mad at Sandra for talking on the phone when she was supposed to be bussing tables. You had to work her tables as well. You are mad at her, but she doesn't think she did anything very bad. How much would each fact affect your anger?

- \_\_\_\_\_ Sandra was talking to her boyfriend and finally convinced him to marry her during that phone call
- \_\_\_\_\_ Sandra knew how long she was talking.
- \_\_\_\_\_ Sandra was talking to someone she was dating casually, but didn't like much
- \_\_\_\_\_ Sandra didn't realize how long she was talking.
- \_\_\_\_\_ There were lots of customers that day
- \_\_\_\_\_ Sandra is twenty-two years old
- \_\_\_\_\_ There were not many customers that day

Imagine that a friend ruined one of your sweaters. It is completely useless now, and you are very upset about it. He doesn't think what he did was very bad. How would each fact affect your anger?

\_\_\_\_\_ He ruined your sweater by using it to stop the bleeding from a gash on his head caused by a car crash.

\_\_\_\_\_ He didn't know it was your sweater when he ruined it.

\_\_\_\_\_ He ruined your sweater by using it to clean his tub.

\_\_\_\_\_ He ruined your sweater on purpose.

\_\_\_\_\_ Your grandmother had knit you the sweater before she died.

\_\_\_\_\_ He didn't ruin your sweater on purpose.

\_\_\_\_\_ You got the sweater at a garage sale for a dollar.

\_\_\_\_\_ He knew it was your sweater when he ruined it.