

Revenge, even though it is not your fault

doi:10.1017/S0140525X12000623

Rongjun Yu

Department of Psychology and Center for Studies of Psychological Application, South China Normal University, Guangzhou, 510631, China.
 rongjun.yu@gmail.com www.rongjunyu.com

Abstract: McCullough et al. argue that revenge has a future-oriented function, that is, to deter future harms by changing other individuals' incentives toward the self. Recent research has shown that people seek revenge even when harms are unintentional. This commentary reports these results and proposes that revenge may also serve to reduce the immediate psychological pain resulting from unfair treatment.

Two factors play a key role in criminal conviction in the common law tradition: a harmful consequence (*actus reus*) and the intent to harm (*mens rea*). Intentions at the time of action influence moral judgment and the subsequent punishment (Young et al. 2010). Individuals who harm others accidentally and unknowingly receive less punishment than those who harm others intentionally. McCullough et al. propose that revenge functions to deter future harms by increasing others' "welfare tradeoff ratios" (WTRs) toward victims. This theory implies that revenge should exist only when harms are intentional because only intentional harms can reveal others' WTRs. On the other hand, unintentional harms are not informative about the harm doer's true WTR and thus should not invite revenge. However, punishment of innocent people is not uncommon in real life situations. Take envy, for example; disliking others' wealth leads people to pay to destroy the envied person's money, even though the envied person is not responsible for the inequality. In laboratory studies, evidence suggests that revenge exists even when harms are unintentional.

Our recent functional magnetic resonance imaging (fMRI) study investigated the behavioral and neural responses to different types of fairness (Yu et al., submitted). In our experiment (see Fig. 1), two participants (strangers) jointly completed a matching task and then they received the outcomes. If their choices were matched, they both received some monetary rewards. Otherwise, they both lost money. However, the exact amounts of money each player could win or lose in each trial were determined by a computer program. Participants received advantageous (more than their partner), disadvantageous (less than their partner), or equal payoffs. Then, they were given the opportunity to alter their partner's payoff at their personal costs. It is a one way punishment, that is, their partners did not have such opportunity to punish. Every fifty pence increase or decrease in the partner's payoff cost participants ten pence.

We found that individuals have strong preferences for fairness in both disadvantageous and advantageous inequality conditions, such that they alter others' payoff toward an equal distribution at a personal financial cost. At the neural level, individuals who spent more money to increase others' payoff had stronger activity in the putamen (the reward region) when they encountered advantageous inequality (Mobbs et al. 2009). Conversely, those who spent more money to reduce others' payoff had stronger activity in the amygdala (the anger region) in response to disadvantageous inequality, suggesting that negative emotions evoke revenge (Scott et al. 1997). Revenge may reduce the immediate psychological harms (e.g., envy and anger) by bringing others down. Our study suggests that accidental harms are enough to elicit immediate negative emotions which may evoke the desire for revenge.

Why do people punish those who are not responsible for the inequality? One possibility is that when an individual is treated unfairly, the induced negative emotion is quite intense. Like physical pain, which makes people punch objects, and frustration,

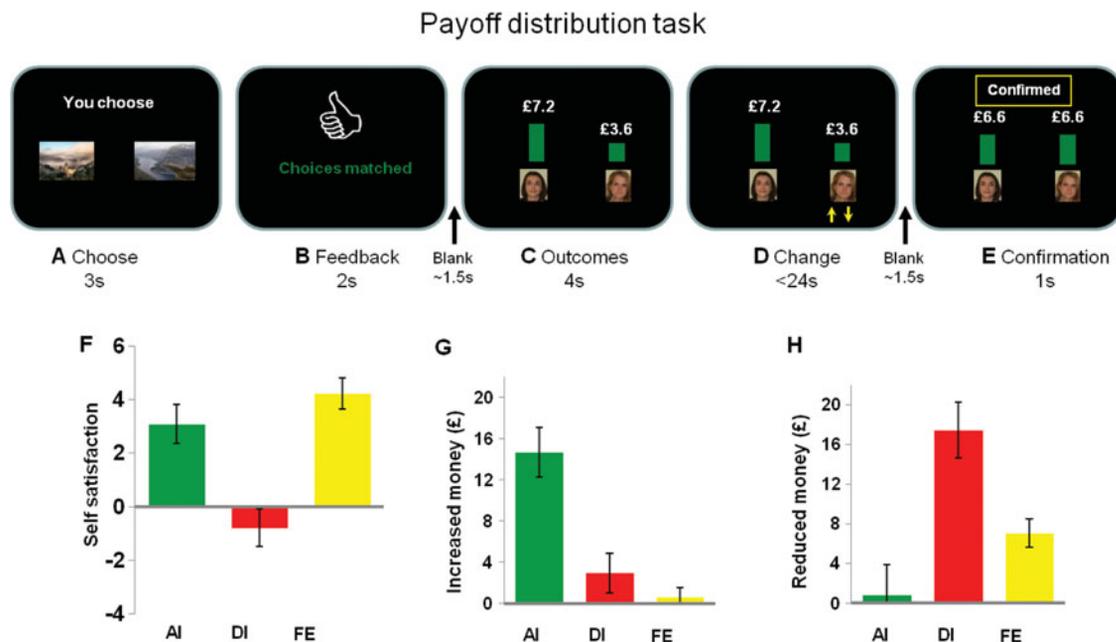


Figure 1 (Yu). **Experimental task design and behavioral results.** (A) In the payoffs distribution task, participants were required to choose one image. (B) After the Choose stage, participants were informed whether their choices were matched or not, and hence, both win or both lose. (C) The outcome for the participants and the outcome for their partners were presented. (D) After the Outcome stage, participants could alter the partner's payoff at their personal costs. (E) Participants pressed a third key when they finished changing. The final payoffs for both players were depicted. (F) The self-reported satisfaction toward outcomes across win and loss trials in advantageous inequality condition (AI), disadvantageous inequality condition (DI), and fair equal condition (FE). (G) The increased money (total money spent to increase other's payoffs) in each condition. (H) Reduced money (total money spent to decrease other's payoffs) in each condition.

which provokes displaced aggression, the social pain resulting from inequality drives people to revenge. The psychological urge to reduce immediate pain is ignored in the target article. From an evolutionary perspective, the sense of unfairness is vital for an individual's survival in social situations and thus revenge may have evolved as an instinctive reaction to unfair treatment. Furthermore, outcomes are easy to evaluate but intentions are difficult to know. Negative outcomes may be enough to elicit revenge motives in the initial stage. Whether to take revenge or to forgive is modulated by attribution of intentions in the latter stage. Previous research shows that forgiveness requires the effort to restrain vengeful impulses (DeWall et al. 2007; 2010), suggesting that revenge is an emotional "hot" system and forgiveness is a rational "cool" process.

Revenge may ultimately hurt the seeker as much as the victim (Dreber et al. 2008). Like a proverb states, "Before you embark on a journey of revenge, dig two graves." In our study, punishment reduces both players' payoffs and participants know that. Even when revenge seekers know such consequences, for example, for individuals who commit crime of passion, they still choose to do so. I argue that revenge is not always future-oriented and may have evolved for other reasons, such as the fairness instinct. It is possible that revenge serves to restore the fairness social norm when individual self-interest has been violated by others. It functions mainly to reduce current emotional harms rather than to deter future harms. This explains why in many situations people seek revenge even when it escalates conflicts rather than moderates them, leading to destructive outcomes for everybody involved.

Authors' Response

doi:10.1017/S0140525X12001513

Putting revenge and forgiveness in an evolutionary context

Michael E. McCullough,^a Robert Kurzban,^{b,c} and Benjamin A. Tabak^d

^aDepartment of Psychology, University of Miami, Coral Gables, FL 33124-0751; ^bDepartment of Psychology, University of Pennsylvania, Philadelphia, PA 19104; ^cEconomic Science Institute, Chapman University, Orange, CA 92866; ^dDepartment of Psychology, University of California–Los Angeles, Los Angeles, CA 90095-1563.

mikem@miami.edu

<http://www.psy.miami.edu/faculty/mmccullough>

kurzban@psych.upenn.edu

<http://www.psych.upenn.edu/~kurzban/>

btabak@psych.ucla.edu

Abstract: In this response, we address eight issues concerning our proposal that human minds contain adaptations for revenge and forgiveness. Specifically, we discuss (a) the inferences that are and are not licensed by patterns of contemporary behavioral data in the context of the adaptationist approach; (b) the theoretical pitfalls of conflating proximate and ultimate causation; (c) the role of development in the production of adaptations; (d) the implications of proposing that the brain's cognitive systems are fundamentally computational in nature; (e) our preferred method for considering the role of individual differences in computational systems; (f) applications of our proposal to understanding conflicts between groups; (g) the possible implications of our views for understanding the operation of contemporary criminal justice systems; and (h) the question of whether people ever "genuinely" forgive.

R1. Introduction

We are grateful to the many scholars who took the time to read and consider our target article. Despite their potential importance to social life, revenge and forgiveness have been, we think, undertheorized (McCullough 2008), and it was our hope that through an adaptationist analysis of behavior and a computational understanding of cognition we might help to stimulate the sorts of research projects in the future that would contribute to a fruitful consilience of the social, behavioral, and life sciences (E. O. Wilson 1998). As **Konečni** points out, the scientific record is full of important empirical results that are relevant to our claims, although inevitably we failed to find all of them. We are thankful for those that commentators such as Konečni have brought to our attention.

The commentators have raised issues of two broad types: first, those that concern the specific claims about revenge and forgiveness that emerged from our approach, and, second, those that concern the meta-theoretical apparatus we put to work in our analysis. We broke responses down further into eight substantive themes. In this response we take the eight themes in turn. In Section R2 we discuss the inferences that are and are not licensed by patterns of contemporary behavioral data in the context of the adaptationist approach. In Section R3 we describe the theoretical pitfalls of conflating proximate and ultimate causation. In Section R4 we clarify our stance on the role of development in the assembly of adaptations. In Section R5 we lay out the implications of proposing that the brain's cognitive systems are fundamentally computational in nature. In Section R6 we describe our approach to considering the role of individual differences in computational systems. In Section R7 we comment on the possible applications of our theorizing to conflicts between groups. In Section R8, we explore the possible implications of our views for understanding the operation of contemporary criminal justice systems. Finally, in Section R9 we consider the question of whether people ever genuinely forgive.

R2. What are the entailments of claiming that cognitive mechanisms for revenge and forgiveness are adaptations with identifiable functions?

Some commentators believe our analysis of revenge and forgiveness leads to implausible hypotheses about the widespread occurrence of revenge in human societies (**Gintis**), that the analysis "sheds very little light on the evolution of these purported cognitive systems" (**Wreha & Racine**), and that adaptations for revenge and forgiveness are unlikely to exist at all (**Holbrook, Fessler, & Gervais [Holbrook et al.]**). **Barclay** is right in pointing out that one major risk of adaptationist analysis is that readers might misperceive functional claims as universal claims. The claim that *the revenge system is an adaptation* emphatically does not entail that *all instances of revenge (or forgiveness) will be adaptive* (Andrews et al. 2002; West-Eberhard 1992; Williams 1966).

Gintis bases his skepticism of our central claims, which he calls "implausible," on (a) evidence from economics experiments indicating that third parties will, under some laboratory conditions, pay costs to punish a stranger who has been stingy or greedy with regard to another stranger;