Third-Party Retaliation and the psychology of deterrence: Mapping the psychological mechanisms that regulate retaliation on behalf of others

Michael McCullough
UNIVERSITY OF MIAMI

09/03/2015
Final Report

DISTRIBUTION A: Distribution approved for public release.
Third-Party Retaliation and the psychology of deterrence: mapping the psychological mechanisms that regulate retaliation on behalf of others

McCullough, Michael E

University of Miami
1320 S. Dixie Highway, Locator Code 2960
Coral Gables, FL 33146

Dr. Benjamin Knott
AF Office of Scientific Research
875 North Randolph Street
Arlington, VA 22203

“Approved for Public Release; Distribution is Unlimited.”

The results of our three-year project can be summarized as follows: People seem quite hesitant to punish strangers who have harmed other strangers. In most of the laboratory experiments we conducted, and in the real world, people retaliate on behalf of themselves and their friends because such harms make them angry (and, in the case of punishment on behalf of friends, because they experience empathy for their friends who have been victimized). In the real world, the welfare trade-off measure appears to be useful for predicting when people will intervene on behalf of a victim; it might turn out to be less useful for predicting third-party punishment under laboratory conditions in which the harms that can be manipulated experimentally, and the kinds of retaliation that can be imposed on others, are necessarily mild for ethical reasons. We conducted six experiments and two non-correlational field studies, and are developing a theoretical paper.

Revenge, retaliation, third-party punishment, welfare interdependence, welfare trade-offs, social psychology.
INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATES COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33615-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report, e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. 61101A.

5d. PROJECT NUMBER. Enter all project numbers as they appear in the report, e.g. 1F665702D1257; ILIR.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.
Final Performance Report

Contract/Grant Title: Third-Party Retaliation and the Psychology of Deterrence: Mapping the Psychological Mechanisms that Regulate Retaliation on Behalf of Others

Principal Investigator: Michael E. McCullough

Contract/Grant Number: FA9550-12-1-1079

Reporting Period: 1 May 2012 – 30 Apr 2015

Project Overview

Traditional theories of deterrence typically consider two types of actors in the midst of conflict. First, there are actors who perceive themselves to be threatened by other individuals or entities, and who try to signal their willingness to impose retaliatory costs on others to deter would-be adversaries. Second, there are the intended recipients of those signals, including (as yet) uninvolved third parties. Signals—demonstrations of force and willingness to use it—are made to discourage adversaries, as well as bystanders, from harming the signaler.

Retaliation is an elementary form of deterrence that has considerable relevance for understanding international conflict. Elsewhere, our research group has proposed that the function of revenge is to deter harmdoers from harming avengers again in the future. To create these deterrent effects, revenge mechanisms motivate avengers to impose retaliatory costs on their harmdoers, which in turn causes harmdoers to revise downward their estimates of the rate of return they can expect from engaging in similar exploitive behaviors against the retaliator in the future. In short, we proposed that retaliation deters aggressors by teaching them that future harms against the avenger will be unprofitable. The same psychological mechanisms that motivate revenge between individuals are very likely involved in strategic decision-making about the use of force at the small group, large group, and inter-state levels. Although our proposal that the function of revenge is to deter future harms may seem intuitive to some readers, we hasten to note that many (if not most) social scientists would disagree. Indeed, most social scientists take very different positions on the functions of revenge, proposing variously that its function is, for example, to re-build threatened self-esteem, to obtain enjoyment or satisfaction, or to balance a moral ledger.

Laboratory social science research on the psychology of retaliatory deterrence has, in keeping with traditional deterrence theory, focused largely on episodes involving only two people: One person (a harmdoer) mistreats a second person (a victim), and the victim then decides whether to retaliate. Largely ignored has been the role of third parties in retaliation.

The work we conducted in this grant-funded project was therefore designed to address the largely unanswered question about the role of third parties in such conflict situations: Why (and under what conditions) do third parties get involved in others’ conflicts, especially when such interventions expose those third parties to potentially steep costs (including, in the military setting, personnel casualties and the destruction of important assets) that otherwise could have been avoided? Our Central Hypothesis was that third-party retaliation is a form of “revenge by proxy” whose function is to deter aggressors from damaging a victim whose welfare is interdependent with the retaliator’s own welfare.

This hypothesis led us to a central prediction: Third-party retaliation will depend on the perceived interdependence between the welfare of the victim and the welfare of the prospective retaliator. Social psychology research demonstrates that in many domains of cognition and behavior, individuals assess the extent to which their welfare is interdependent with the welfare of others. We posit that perceived welfare interdependence regulates how cognitive and physiological resources are allocated, with more resources deployed to process and respond to information about people with whom we are interdependent. As noted above, there is good reason to believe that the psychological systems that support the computation of welfare interdependence and the motivation of behavior based on those computations in individual relations are also relevant for understanding decisions about the use of force in military conflicts.

To investigate the central hypothesis in detail, this project was designed to address the following four objectives:

1. Model the computational process by which actors assess the interdependence of their welfare with the welfare of another actor;
(2) Examine three discrete social-psychological conditions (i.e., friendship, recent generosity, and upcoming opportunities for cooperation) that the human mind might use as cues for estimating interdependence of one’s welfare with that of another individual;

(3) Examine the role of perceived welfare interdependence as a mediator of the effects of friendship, recent generosity, and upcoming opportunities for cooperation on third-party retaliation; and

(4) Identify the role of empathy for victims and anger toward aggressors as intermediate psychological processes that motivate third-party retaliation.

Summary of Specific Findings

Below, we summarize the findings that have emerged from this project so far with regard to the four specific objections enumerated above. Relevant references to papers are denoted with numerals that refer to specific items in the reference list at the end of this document.

(1) Modeling the computational process by which actors assess the interdependence of their welfare with the welfare of another actor.

In Paper 2, we found that subjects were reticent to pay costs in order to retaliate against strangers who had harmed another stranger, even though they were relatively willing to retaliate against a stranger who had harmed them directly. This initial finding led us to anticipate that people computed their welfare interdependence with others on the basis of cues such as friendship, cooperation, and opportunities for future cooperation. Paper 4 is our attempt to situate third-party anger and third-party retaliation within a general theory of the function of anger (which is, we posit, the emotion that motivates deterrence behavior). In Paper 4, we offer an explicit formulation for how welfare interdependence might cause anger toward individuals who have harmed either oneself or another individual.

Throughout several of our data collection efforts, we used a recently developed self-report measure for estimating the extent to which one individual would be willing to reduce his own welfare in order to raise the welfare of another individual. This welfare trade-off measure is simple and quick to administer. In Papers 5, 6, 7, 10, and 12, we used this measure to evaluate whether people’s willingness to retaliate on others’ behalf in the laboratory appeared to be mediated by their welfare trade-off ratios for those victims. The data for Papers 7, 10, and 12 have been analyzed, and to our surprise, we found fairly little evidence in support of our hypothesis. In contrast, Paper 8—which consists of correlational field data not only from the US but also from Japan—indicates that people do indeed base their willingness to intervene on behalf of others on their welfare trade-off ratios for those individuals: the more you value someone’s welfare (out in the real world) the more likely it is that you will intervene on their behalf when another individual is imposing a harm upon them. After we have more thoroughly analyzed data for Papers 5 and 6, we should have a better idea of whether the welfare trade-off ratio measure can be productively used in laboratory settings to measure changes in people’s welfare evaluations of strangers with whom they have had only minimal interaction.

(2) Examine three discrete social-psychological conditions (i.e., friendship, recent generosity, and upcoming opportunities for cooperation) that the human mind might use as cues for estimating interdependence of one’s welfare with that of another individual.

Papers 5, 7, 10, and 12 were designed to evaluate the effects of friendship and recent generosity and upcoming opportunities for cooperation on people’s welfare trade-off ratios for victims and their subsequent willingness to punish on those victims’ behalf. In Paper 7, we found evidence that people are indeed willing to punish harmdoers who have harmed their friends, but that they are unwilling to punish on behalf of complete strangers—even when they are capable of punishing at no cost whatsoever to themselves. In contrast, the results of Papers 10, 11, and 12 suggest that people are not willing to punish on behalf of themselves, friends, or strangers, when the harm committed was an ungenerous allocation in a modified Dictator Game.

(3) Examine the role of perceived welfare interdependence as a mediator of the effects of friendship, recent generosity, and upcoming opportunities for cooperation on third-party retaliation.
We have collected ample data for addressing this objective. So far, the only data sets in which we have done so are the data sets for Papers 7, 10, 11, and 12. The data were not supportive of this prediction. Papers 10, 11, and 12 show that although welfare interdependence was higher when the victim was a friend rather than a stranger, there was no relationship between friendship and third-party retaliation. However, the data set for Paper 5 also is relevant to this question and will be analyzed shortly.

(4) Identify the role of empathy for victims and anger toward aggressors as intermediate psychological processes that motivate third-party retaliation.

Several of our papers address this objective. Paper 4 addresses the issue theoretically, and Papers 5, 6, 8, 9, 10, 11, and 12 address the issue empirically. Papers 5 and 9 both show that people are willing to retaliate when they are harmed directly or when a friend is harmed precisely because they are angered by such harms. In contrast, people do not reliably experience anger when they witness one stranger harm another stranger (Paper 9), so anger is apparently incapable of explaining the small amount of third-party retaliation we see in that experiment. Likewise, we found that third parties did not retaliate on strangers' behalf in the data set for Paper 5, so there was no third-party retaliation to be explained on the basis of emotions such as anger and empathy. In Papers 10 and 11, there was no effect of relationship with the victim on third-party anger, nor were vengeful third-parties angrier than those who chose not to punish the harmdoer. Note that this was expected for Paper 11, in which the experimenters explicitly made the unfair allocations, and thus the would-be harmdoers would not be the appropriate targets for anger. Third parties in Paper 11 were, however, more envious of the recipients of the larger portion of the allocation. Peculiarly, the results from Paper 12 suggest that people experience more anger when observing a stranger receiving an unfair allocation than a friend experiencing the same harm. Additionally, those who chose to reward harmdoers reported the most anger. These may have been spurious findings, but future research should aim to elucidate these issues.

So far, then, the results of our project can be summarized as follows. People seem quite hesitant to punish strangers who have harmed other strangers. This lack of third-party punishment was readily observed in our laboratory experiments and was also made plain by the field studies that we conducted as part of Paper 8. In most of the laboratory experiments (except Papers 10, 11, and 12) and in the real world, people retaliate on behalf of themselves and their friends because such harms make them angry (and, in the case of punishment on behalf of friends, because they experience empathy for their friends who have been victimized). In the real world, the welfare trade-off measure appears to be useful for predicting when people will intervene on behalf of a victim; it might turn out to be less useful for predicting third-party punishment under laboratory conditions in which the harms that can be manipulated experimentally, and the kinds of retaliation that can be imposed on others, are necessarily mild for ethical reasons.

Potential for Translation to Military Applications

Several findings from this project have potential translational value. First, we have demonstrated the potential utility of conceptualizing alliances and friendships in terms of welfare interdependence. We administered the welfare trade-off measure used in these studies to thousands of subjects and have found it to be easy to understand and relatively quick to complete. On the basis of our field studies (Paper 8), welfare trade-off measurements appear to predict people’s willingness to intervene aggressively in the disputes of other individuals (people intervene on behalf of people with whom they perceive their welfare to be interdependent). It seems likely that this measure could be used on the ground by war fighters to quickly assess various agents’ willingness to engage in third-party intervention on behalf of other entities at many levels of complexity. For example, the measure could be used sociometrically to map individual alliances within a neighborhood, or within a group of insurgents, or even to estimate the strengths of alliances among global players (perhaps using more objective measures such as trading arrangements and treaty formation). Furthermore, the measure could potentially be adapted to tracking the flow of real-world resources between individuals to gain objective measures of welfare interdependence.

Additionally, findings from Papers 1 and 3—broadly, that harsh conditions in one’s local environment during childhood and adolescence lead to increased social distrust and willingness to violently retaliate later in life—
suggest the potential for identifying hotspots in communities for producing youth susceptible to recruitment by terrorist groups and other radical organizations. In the future, we look to expand this work by developing objective measures of environmental harshness via geographic information systems (GIS).

**Performance Metrics**

**Academic**

The project included six laboratory experiments, four field studies (two of which were directly funded by AFOSR plus two others completed independently by our research partners in Japan and Indonesia). We anticipate no fewer than seven journal articles to emerge from this project over the coming several years as we analyze data and write them up. We have found that our work on punishment is attractive at some of the best journals in Psychology and the General Sciences, so we are optimistic that many of these papers will go on to influence many scientists’ thinking about third-party intervention in agonistic contexts. We will be making the data sets from this project publicly available (after the data have been completely de-identified) so that future researchers might use them to ask additional questions and publish additional papers based on the work we accomplished here.

**Training**

Four graduate students, 23 undergraduate students, two paid research assistants, and two high school students were trained as a direct result of this grant. One graduate student, Eric Pedersen, completed his Ph.D. and is beginning a postdoctoral fellowship at the University of Miami. A second graduate student, Daniel Forster, completed his master’s thesis and is beginning work on his dissertation. A third graduate student, William McAuliffe, began work in our lab as an undergraduate research assistant for the grant and is now beginning work on his master’s thesis after completing his first year of graduate school. A fourth graduate student, Fatima Aboul-Seoud, completed her first year of her Ph.D. program at the University of Pennsylvania and successfully defended her master’s thesis, in which she wrote up three of the studies conducted for this grant.

**Publications**

**Publications Relevant to the Grant that Were Published During the Grant Period:**


**Papers in Progress that are Directly Attributable to this Grant:**


(11) Aboul-Seoud, F., McCullough, M. E., & Kurzban, R. (in prep-b). Is punishment due to a desire for revenge or an aversion to inequality?

(12) Aboul-Seoud, F., McCullough, M. E., & Kurzban, R. (in prep-c). Do observers fail to punish on behalf of unfairly treated stranger-recipients because they believe that those stranger-recipients are capable of punishing for themselves?

Presentations


1. Report Type
   Final Report

Primary Contact E-mail
Contact email if there is a problem with the report.
   mikem@miami.edu

Primary Contact Phone Number
Contact phone number if there is a problem with the report
   305-342-5277

Organization / Institution name
   University of Miami

Grant/Contract Title
The full title of the funded effort.
   Third-Party Retaliation and the psychology of deterrence: mapping the psychological mechanisms that regulate retaliation on behalf of others

Grant/Contract Number
AFOSR assigned control number. It must begin with “FA9550” or “F49620” or “FA2386”.
   FA9550-12-1-0179

Principal Investigator Name
The full name of the principal investigator on the grant or contract.
   Michael E. McCullough

Program Manager
The AFOSR Program Manager currently assigned to the award
   Dr. Benjamin Knott

Reporting Period Start Date
   05/01/2012

Reporting Period End Date
   04/30/2015

Abstract
Over the past year, we completed the remaining experiments at the Miami and Penn sites. Over the project period, six experiments were conducted in total (3 at Miami; 3 at Penn). In addition, we completed two non-correlational field studies in which we evaluated people’s likelihood of being angered and retaliating on behalf of others in real-life (retrospectively recalled) conflict events. These studies were extended and replicated (not using AFOSR support) by a collaborator in Japan (who has finished collecting and analyzing data) and Indonesia (who has finished collecting data, but has not fully analyzed it). In addition, we worked with a collaborator in Venezuela to conduct a Venezuelan replication in Fall 2015. Finally, we are developing a theoretical paper to show how estimates of welfare interdependence can explain why bystanders become angry and vindictive when they observe that one individual has been harmed by another individual. The results of our three-year project can be summarized as follows: People seem quite hesitant to punish strangers who have harmed other strangers. In most of the laboratory experiments we conducted, and in the real world, people retaliate on behalf of themselves and their friends because such harms make them angry (and, in the case of punishment on behalf of friends, because they experience empathy for their friends who have been victimized). In the real world, the welfare trade-off measure appears to be useful for predicting when people will intervene on behalf of a victim; it might turn out to be
less useful for predicting third-party punishment under laboratory conditions in which the harms that can be manipulated experimentally, and the kinds of retaliation that can be imposed on others, are necessarily mild for ethical reasons. Possible military application: Evaluations of welfare interdependence among actors—using self-report methods or other methods—to assess alliances, friendships, and other forms of relationships that would indicate a willingness to pay costs in order to benefit other actors (or their causes).

Distribution Statement
This is block 12 on the SF298 form.

Distribution A - Approved for Public Release

Explanation for Distribution Statement
If this is not approved for public release, please provide a short explanation. E.g., contains proprietary information.

SF298 Form
Please attach your SF298 form. A blank SF298 can be found here. Please do not password protect or secure the PDF. The maximum file size for an SF298 is 50MB.

Final SF298 _Submitted.pdf

Upload the Report Document. File must be a PDF. Please do not password protect or secure the PDF. The maximum file size for the Report Document is 50MB.

FINAL PERFORMANCE REPORT_15_7_31.pdf

Upload a Report Document, if any. The maximum file size for the Report Document is 50MB.

Archival Publications (published) during reporting period:
None published yet, though several are in preparation.

Changes in research objectives (if any):
None.

Change in AFOSR Program Manager, if any:
Dr. Benjamin Knott continued as the Program Manager.

Extensions granted or milestones slipped, if any:
None.

AFOSR LRIR Number

LRIR Title

Reporting Period

Laboratory Task Manager

Program Officer

Research Objectives

Technical Summary

Funding Summary by Cost Category (by FY, $K)

<table>
<thead>
<tr>
<th></th>
<th>Starting FY</th>
<th>FY+1</th>
<th>FY+2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipment/Facilities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supplies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Report Document
Report Document - Text Analysis
Report Document - Text Analysis

DISTRIBUTION A: Distribution approved for public release.
Appendix Documents

2. Thank You

E-mail user

Jul 30, 2015 13:33:54 Success: Email Sent to: mikem@miami.edu