

Humans reciprocate intentional harm by discriminating against group peers

David Hugh-Jones^{*} Itay Ron[†] Ro'i Zultan[‡]

Abstract

The evolution of human intergroup conflict is a social science puzzle. Motivated by cycles of intergroup revenge in real-world conflicts, we experimentally test the hypothesis that humans practice group-based reciprocity: if someone harms or helps them, they harm or help other members of that person's group. Subjects played a trust game, then allocated money between other people. Senders whose partners returned more in the trust game gave more to that partner's group members. The effect was about half as large as the effect of direct reciprocity. Receivers' allocations to group members were not affected by their partners' play in the trust game, suggesting that group reciprocity was only triggered when the partner's intentions were unequivocal. We show conditions under which group reciprocity can evolve, and discuss its place in conflict among early humans.

Keywords: Upstream reciprocity, group identity, intergroup conflict.

Word count: 3129

^{*}School of Economics, University of East Anglia. E-mail: D.Hugh-Jones@uea.ac.uk.

[†]E-mail: itayron@gmail.com

[‡]Corresponding author, Department of Economics, Ben-Gurion University of the Negev. E-mail: zultan@bgu.ac.il.