Journal of Experimental Social Psychology 73 (2017) 76–81

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Case Report

Deliberation erodes cooperative behavior — Even towards competitive out-groups, even when using a control condition, and even when eliminating selection bias

Jim A.C. Everett[a,b,*], Zach Ingbretsen[a], Fiery Cushman[a], Mina Cikara[a]

[a] Department of Psychology, Harvard University, Cambridge, MA, USA
[b] Department of Experimental Psychology, Oxford University, Oxford, UK

ABSTRACT

By many accounts cooperation appears to be a default strategy in social interaction. There are, however, several documented instances in which reflexive responding favors aggressive behaviors: for example, interactions with out-group members. We conduct a rigorous test of potential boundary conditions of intuitive prosociality by looking at whether intuition favors cooperation even towards competitive out-group members, and even in losses frames. Moreover, we address three major methodological limitations of previous research in this area: a lack of an unconstrained control condition; non-compliance with time manipulations leading to high rates of exclusions and thus a selection bias; and non-comprehension of the structure of the game. Even after eliminating participant selection bias and non-comprehension, we find that deliberation decreases cooperation: even in competitive contexts towards out-groups and even in a losses frame, though the differences in cooperation between groups was consistent across conditions. People may be intuitive cooperators, but they are not intuitively impartial.

1. Introduction

A recent body of research suggests that, far from requiring effortful control, behaving prosocially arises from "processes that are intuitive, reflexive, and even automatic" (Zaki & Mitchell, 2013, p. 466). Several studies find that people tend to make prosocial decisions in economic games more quickly than selfish ones, and time-pressure increases the incidence of prosocial behavior (for an overview and meta-analysis see Rand, 2016). Time delay reduces helping in a 'dropped-glove' field study (Artavia-Mora, Bedi, & Rieger, 2017) and even risking one's own life to save another seems driven primarily by intuitive processes (Rand & Epstein, 2014). To explain this, the *Social Heuristics Hypothesis* (SHH: Rand, 2016; Rand et al., 2014) posits that the social strategies typically successful in daily life (e.g., cooperation) become automatized as default responses, and that deliberation can override these defaults to modify behavior. Indeed, a formal analysis of evolutionary dynamics indicates that deliberation can only serve to undermine costly cooperation and not promote it (Bear, Kagan, & Rand, 2017; Bear & Rand, 2016). While meta-analytic work has provided strong support for the claim that manipulating reliance on intuition through time pressure encourages prosociality (Rand, 2016; Rand et al., 2014), a recent Registered Replication Report (RRR) by Bouwmeester et al. (2017) finds only mixed support. Our aim is to consider two challenges to the idea that intuition favors cooperation—one methodological and one theoretical—and then to provide new experimental evidence.

1.1. Boundary conditions on intuitive cooperation: intergroup bias and decision framing

One of the most enduring findings in social psychology is intergroup bias: the powerful tendency to evaluate and treat in-group members more favorably than out-group members (Hewstone, Rubin, & Willis, 2002). How does intuitive cooperation play out in intergroup contexts? The SSH posits that cooperative strategies that are typically advantageous in daily life with repeated trustworthy interaction partners become internalized as intuitions and get overgeneralized to less typical settings. Given this, we might reason that social heuristics, precisely because they are overgeneralized, will apply even to out-group members and it is deliberation that 'corrects' this process and causes increased intergroup bias. Supporting this, for example, is an experiment

* Corresponding author.
E-mail address: jim.ac.everett@gmail.com (J.A.C. Everett).

http://dx.doi.org/10.1016/j.jesp.2017.06.014
Received 26 February 2017; Received in revised form 20 June 2017; Accepted 21 June 2017
Available online 29 June 2017

employing a Public Goods Game with minimal groups; conceptual priming of intuition leads to equivalent contributions to both in-group and out-group members, whereas priming reflection leads to pronounced in-group favoritism (Ma, Liu, Rand, Heatherton, & Han, 2015).

Yet the idea that intuition uniformly favors cooperation even with out-group members is at odds with much past evidence suggesting that the preference for 'us' over 'them' is at least partially rooted in implicit, unconscious, or automatic processes. Preference for in-group members appears early in development (Dunham, Baron, & Banaji, 2008), and group memberships appear to generate bias at the earliest stages of perceptual and emotional processes (for a review see Cikara & Van Bavel, 2014). Intergroup bias often occurs outside the realm of conscious awareness and, for this reason, can be remarkably difficult to control or change permanently (Greenwald, McGhee, & Schwartz, 1998). For example, a recent test of 9 interventions that reduce implicit prejudice in the IAT found that none lasted more than one day (Lai et al., 2016). There is also evidence for intuitive aggression towards out-groups - for example, in the Shooter task where participants are required to "shoot" armed targets and to "not shoot" unarmed targets (Correll, Park, Judd, & Wittenbrink, 2002). Participants under time pressure are more likely to mistakenly shoot other-race targets than same-race targets (Correll et al., 2002) and are more likely to mistakenly shoot minimal out-group members than in-group members (Miller, Zielaskowski, & Plant, 2012).

Looking specifically at prosocial behavior, several convergent findings indicate that intuition encourages in-group favoritism rather than impartial cooperation. Experiments using the Intergroup Prisoner's Dilemma–Maximizing Differences Game (Halevy, Bornstein, & Sagiv, 2008) find that parochial altruism - contributions to a pool that both benefits the in-group and simultaneously hurts the out-group - emerges especially among individuals who were cognitively taxed by completing a Stroop interference task (De Dreu, Dussel, & Velden, 2015). Another series of experiments using a Prisoner's Dilemma with real political groups (Obama vs. Romney supporters) and a time pressure manipulation demonstrate that intuition increases cooperation to both in-group and out-group members, but that the difference between groups is maintained relative to an enforced time delay condition (Rand, Newman, & Wurzbacher, 2015). Therefore, while at least one study indicates that intuition reduces in-group favoritism (Ma et al., 2015), two others indicate that intuition increases in-group favoritism (De Dreu et al., 2015; Ten Velden, Daughters, & De Dreu, 2017), and a fourth study indicates that intuition increases prosocial behavior but neither increases nor decreases in-group favoritism (Rand et al., 2015).

One possibility for this inconsistency is that existing studies using time pressure have typically contrasted a manipulation intended to promote intuition against a manipulation intended to promote reflection, but do not assess these manipulations against an unconstrained control condition (e.g. Capraro, Jordan, & Rand, 2014; Cone & Rand, 2014; Rand et al., 2014; Verkoeijen & Bouwmeester, 2014). Such a design cannot distinguish evidence that time pressure increases cooperation from evidence that time constraint reduces cooperation; as such, discrepancies across the intergroup studies might reflect the relative efficacy of the intuitive versus deliberative manipulations included in each.

Our second theoretical contribution is to explore the consequences of framing a social dilemma in terms of losses versus gains. Motivated in part by prior research on active hostility in intergroup contexts (e.g. Correll et al., 2002; Sherif, 1966), we speculated that defection against out-group members might be the intuitive response especially in a social dilemma framed in terms of losses (i.e., defection imposes a cost) rather than gains (i.e., defection withholds a benefit). Prior research on framing in social dilemmas is inconsistent, and has varyingly shown loss frames to increase cooperation (e.g. Experiment 3 in Komorita & Carnevale, 1992), reduce cooperation (e.g. Brewer & Kramer, 1986), or have no effect (e.g. de Heus, Hoogervorst, & van Dijk, 2010). This may occur because decision frames have divergent effects based on an individual's prior motives, such that prosocial people become more cooperative, and 'individualists' less cooperative (De Dreu & McCusker, 1997). Because people's prior motives are different with regards to in-group and out-group members, we might therefore expect loss frames to encourage defection towards out-groups (because without a prior motive to be cooperative, loss frames make people more selfish) and cooperation towards in-groups (because loss frames enhance the existing motive to help the in-group member).

1.2. Methodological challenges for intuitive cooperation

In addition to these theoretical concerns, recent research has challenged intuitive cooperation findings on methodological grounds (Tinghög et al., 2013). This critique focuses especially on the use of time pressure versus time delay to manipulate the balance of automatic versus controlled inputs into cooperation decisions, specifically citing high levels of participant exclusions and selection bias (Bouwmeester et al., 2017; Tinghög et al., 2013). The original studies excluded participants who failed to make their decisions within the response window. These exclusion rates are typically very high and, more importantly, asymmetric across conditions. For example, in Rand, Greene, and Nowak (2012), 48% of participants failed to make their decisions under time pressure in Study 6, and 46% failed in Study 7 relative to 19% and 10% in the time delay conditions, respectively. While by some accounts these effects hold when including non-compliant participants (Rand, Greene, & Nowak, 2013; Rand et al., 2014), these exclusion practices introduce the possibility of selection bias. Eliminating participants who are too slow from the time pressure condition and not from the time delay condition disrupts random assignment to condition. Therefore, the observed difference in cooperation could be driven by systematic differences between the participants rather than the manipulation. For example, Tinghög et al. (2013) could not successfully replicate Rand et al.'s (2012) results without exclusions and therefore conclude the original findings were "an artefact of excluding the about 50% of subjects who failed to respond on time" (p.427). Consistent with this possibility, a recent pre-registered multi-site replication study (Bouwmeester et al., 2017) reported that two-thirds of participants failed to make decisions within the allotted time and that the effect of time pressure on cooperation was only present when excluding such individuals. Of course, this data is also consistent with the possibility that individuals who fail to conform to the time-pressure treatment therefore show no effect of that treatment.

A similar problem arises with the use of comprehension checks, where large numbers of participants fail to correctly answer comprehension questions about the structure of the game after they have played it. Across the studies reported by Rand et al. (2012), comprehension checks were implemented after the game had been played to avoid suggesting a deliberative mindset to participants prior to decision making. And indeed, as Rand and colleagues report in a supplementary study, participants who complete comprehension questions before making their decision choose to contribute significantly less than those who complete the comprehension questions afterward (Rand et al., 2012). However, in practice this means that participants may play the game while not understanding it. For example, 32% of participants in Study 1 of Rand et al. (2015) failed one or both comprehension checks and yet were included in the final analysis. While the effects of time pressure were robust to controlling for comprehension, this remains a potentially problematic aspect of the dominant methodological design.

Given these concerns (and the lack of the control condition raised in the previous section), we designed a procedure that accomplishes three key methodological goals. First, we drastically reduced exclusion rates due to the response window, achieving an exclusion rate of just 2% of participants enrolled in the study. We accomplished this by providing participants with extensive comprehension training prior to the task employing alternative payoff matrices. This prepared them to quickly