Do infants detect indirect reciprocity?

Marek Meristo [a],[*], Luca Surian [b],[*]

[a] *Department of Psychology, University of Gothenburg, Box 500, 405 30 Göteborg, Sweden*
[b] *Department of Psychology and Cognitive Sciences, University of Trento, Corso Bettini, 31, 38068 Rovereto, Trento, Italy*

ABSTRACT

In social interactions involving indirect reciprocity, agent A acts prosocially towards B and this prompts C to act prosocially towards A. This happens because A's actions enhanced its reputation in the eyes of third parties. Indirect reciprocity may have been of central importance in the evolution of morality as one of the major mechanisms leading to the selection of helping and fair attitudes. Here we show that 10-month-old infants expect third parties to act positively towards fair donors who have distributed attractive resources equally between two recipients, rather than toward unfair donors who made unequal distributions. Infants' responses were dependent on the reciprocator's perceptual exposure to previous relevant events: they expected the reciprocator to reward the fair donor only when it had seen the distributive actions performed by the donors. We propose that infants were able to generate evaluations of agents that were based on the fairness of their distributive actions and to generate expectations about the social preferences of informed third parties.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

A key aspect in human social skills is the ability to cooperate among genetically unrelated individuals. Despite the potentially conflicting interests, in terms of reproductive fitness, people often choose to behave in a way that benefits one another (Joyce, 2007; Nowak, 2012). By definition, altruistic behavior consists of actions that are intended to benefit another individual while being costly for the actor. The reward for the individual performing the behavior is anything but clear (Joyce, 2007; Nowak & Sigmund, 2005). Within an evolutionary framework, it has been difficult to explain how the ability to engage in such helping behavior towards non-kin could arise, as it seems to contradict a fundamental tenet of natural selection theory: morphological or behavioral traits are selected for only if they are advantageous for the individuals who display

them, that is if they increase their reproductive fitness (Alexander, 1987; Krebs, 2006; Trivers, 1971). In humans, a similar problem emerges when preferences for fairness are concerned, given that such preferences often appear to conflict with the self-interest.

To solve this puzzle, evolutionary biologists have proposed a number of processes that, in genetically 'selfish' terms, may lead to the emergence of altruism and fair behaviors. Indirect reciprocity is one of the most important of such processes (Alexander, 1987; Nowak, 2006; Sober & Wilson, 1998). In social interactions that involve indirect reciprocity, an individual who helps another can potentially gain benefits from her actions by enhancing her reputation in the eyes of interested audiences (Alexander, 1987; Fehr, 2004; Nowak, 2006; Nowak, 2012; Wedekind & Milinski, 2000). Schematically, in the typical case of indirect reciprocity, A helps B, and therefore receives help from C (Fig. 1).

In an attempt to identify the evolutionary roots of altruism, Alexander (1987) formulated a model of 'selection by reputation' where the main premises include potential advantages in the context of marriage choices for the

* Corresponding authors. Tel.: +46 31 7861664; fax: +46 31 7864628 (M. Meristo).

E-mail addresses: marek.meristo@gu.se (M. Meristo), luca.surian@unitn.it (L. Surian).
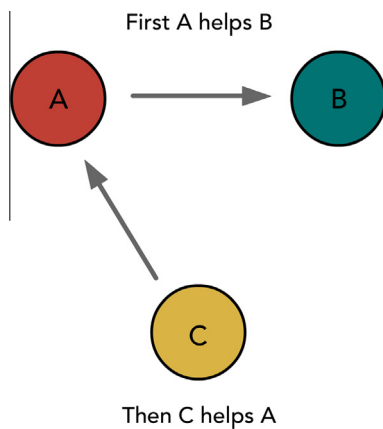
Fig. 1. Schematic illustration of interaction according to the principle of indirect reciprocity. First individual A helps B, and then receives help from C (from Nowak and Sigmund (2005)).

altruistically generous and fair individual in the practices of hunter–gatherer societies. Other forms of compensation for altruists, through reputation building, may include better opportunities for future collaboration in a variety of social partnership contexts, as well as between-group advantages for bands with more altruists. In a similar manner, indirect reciprocity may include negative reputation and its consequences of ostracism or social shunning for group members who do not comply with the rules of conduct or do not reciprocate adequately (Alexander, 1987; Rockenbach & Milinski, 2006; Ule, Schram, Riedl, & Cason, 2009). In indirect reciprocity processes therefore, the costs of extra-familial generosity generally are compensated for by enhanced fitness, showing that altruism towards genetically unrelated individuals is compatible with natural selection. Additionally, the need to keep track of all group members' social histories, in a context of indirect reciprocity, may have been a crucial evolutionary force driving language skills and advanced social cognition (Nowak & Sigmund, 2005).

An important source of evidence concerning the origins and ontogenesis of indirect reciprocity in humans comes from research with infants and young children. While early works on reciprocal altruism have mainly focused on the processes of direct reciprocity (Berndt, 1979; Dreman & Greenbaum, 1973; Keil, 1986; Levitt, Weber, Clark, & McDonnell, 1985; see also Dunfield & Kuhlmeier, 2010), studies on the development of explicit moral reasoning suggest that preschoolers are not capable of verbally reasoning about indirect reciprocity principle (Kenward & Dahl, 2011). However, when preschoolers are asked explicitly to evaluate social actions, they show that they are able to take into account norm violations against third parties (Catron & Masters, 1993; Ingram & Bering, 2010).

Two recent studies have reported that preschoolers act coherently with the indirect reciprocity principle when they are told to distribute scarce resources (Kenward & Dahl, 2011; Olson & Spelke, 2008). Olson and Spelke (2008) asked children to help a puppet to distribute some resources between two pairs of individuals after being told

stories about their previous sharing behavior. When the number of resources was less than the number of recipients, 3.5-year-olds took into account whether the recipients had previously shown generous behavior toward third parties. They preferred however an equal distribution when the number of resources equaled the number of recipients. Kenward and Dahl (2011) showed preschoolers a puppet that was trying to climb a hill. After several unsuccessful attempts, the puppet was either aided by another puppet (the helper), or pushed down by a third puppet (the hinderer). Finally, the children were asked to distribute biscuits between the helper and hinderer. When they had three biscuits to give away, 4.5-year-olds gave more to the helper than to the hinderer, while they preferred an equal distribution when they had eight biscuits. Thus, previous research demonstrated that preschoolers, despite their limitations in the reasoned application of principle of indirect reciprocity, do act in accordance with such principle in some contexts that require them to distribute resources. In order to test whether even preverbal infants are able to make nuanced third party social evaluations and detect indirect reciprocation, we need procedures that rely on non-verbal, spontaneous responses.

To our knowledge, only one previous study has provided such evidence. Hamlin, Wynn, Bloom, and Mahajan (2011) showed 8-month-old infants short events involving puppets performing either helping or hindering actions. They then saw two new agents, the 'giver' and the 'taker', acting prosocially (i.e. giving a ball) or antisocially (i.e. taking away the ball) towards the helping or the hindering puppet. When infants were encouraged to select one of the two new agents they consistently chose those who acted prosocially towards helpers and antisocially towards hinderers. That is, infants preferred puppets who acted consistently with the principle of indirect reciprocity.

Note, however, that in Hamlin et al. (2011) study, it was not clear whether the reciprocating agents knew or did not know the relevant helping/hindering behaviors performed by the main characters. Therefore, in such a context, infants' preferences could reveal just their sensitivity to a Simple Indirect Reciprocity Rule (SIRR) centered on the coherence between the act performed by A and the act performed by C towards A:

– SIRR: *If A helps B, then expect C to help A.*

If this was the case, it would be wrong to conclude that they followed a more Complex Indirect Reciprocity Rule (CIRR), namely a rule that also takes into account the relevant epistemic states of the reciprocator:

– CIRR: *If A helps B, and C knows it, then expect C to help A.*

To be able to reason following SIRR, one only needs to track the congruency between the valence on the actions performed and received by A. By contrast, to follow CIRR, one also needs to track the reciprocator's knowledge of events that are relevant to evaluate A, which is essential in the model of indirect reciprocity. In Hamlin et al. (2011) study, we can assume that antisocial characters