Cognition 126 (2013) 149–155

Contents lists available at SciVerse ScienceDirect

Cognition

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

A non-mentalistic cause-based heuristic in human social evaluations

Marine Buon [a,*], Pierre Jacob [a,b], Elsa Loissel [c], Emmanuel Dupoux [a]

[a] Laboratoire de Sciences Cognitives et de Psycholinguistique, Ecole Normale Superieure, Ecole des Hautes Etudes en Sciences Sociales, Centre Nationale Superieure, Paris, France
[b] Institut Jean Nicod, Ecole Normale Superieure, Ecole des Hautes Etudes en Sciences Sociales, Centre Nationale Superieure, Paris, France
[c] Ecole Normale Superieure, Lyon, France

ARTICLE INFO

Article history:
Received 17 November 2011
Revised 3 September 2012
Accepted 3 September 2012
Available online 21 November 2012

Keywords:
Moral psychology
Dual models
Executive resources
Theory of mind
Causation

ABSTRACT

In situations where an agent unintentionally causes harm to a victim, the agent's (harmless) intention typically carries *more* weight than his/her (harmful) causal role. Therefore, healthy adults typically judge leniently agents responsible for an accident. Using animated cartoons, we show, however, that in the presence of a difficult concurrent task, this result is reversed: the agent's harmless intention is given *less* weight than her harmful causal role, inducing participants to judge harshly the accidental agent. This was found even though cognitive load did not selectively impair the detection of intentions over causal roles. Not only is this finding evidence that the social/moral evaluation system relies on two dissociable components, but it also demonstrates that these components are asymmetrical, the causal component being more intuitive than the intentional component, and the full integration of the two requiring central cognitive resources.

1. Introduction

If someone accidentally steps on your shoes in the street, your immediate response may be to blame him or her. However, after the initial inflow of emotions, you may revise your evaluation and take into account his or her intention. Most psychological models of moral cognition claim that adults give a primary role to the agent's intention to harm when performing moral judgment (Cushman, 2008; Piaget, 1965/1932). Recently, however, researchers have come to recognize that human moral competence is not a unitary system, but rather a collection of heterogeneous components running concurrently, some of which implicate fast emotional responses (Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001), frugal heuristics (Sunstein, 2005) unconscious computations (Cushman, Young, & Hauser, 2006), or more deliberative processes (Cushman et al., 2006; Dupoux & Jacob, 2007; Greene et al., 2008). Examining the effects of the scarcity of cognitive resources on moral judgments is therefore a useful tool for uncovering the cognitive architecture underlying human moral competence.

The case of accidental harm is particularly interesting to study because it requires the resolution of a conflict between the agent's harmful causal role (the victim is harmed) and his/her harmless intention (the agent did not want to harm) (Young, Cushman, Hauser, & Saxe, 2007). Furthermore, the two terms of the conflict may rely on distinct cognitive systems (Cushman, 2008). On the one hand, representing the agent's causal role requires assessing his/her action and the amount of harm endured by the victim. Both computations can be achieved by relatively shallow heuristics: while spatiotemporal correlations help infering causal structure (Michotte, 1946/1963), distress cues and/or emotional contagion help computing the amount of harm (Blair, 1995; de Vignemont & Jacob, 2012). On the other hand, the content of the agent's *intention* must be inferred from prior mentalistic knowledge of

* Corresponding author: Laboratoire de Sciences Cognitives et de Psycholinguistique, Pavillon Jardin, 29 rue d'Ulm, 75005 Paris. Tel.: +33(0)1 44 32 26 27; fax: +33(0)1 44 32 26 30.
E-mail address: marinebuon@gmail.com (M. Buon).

0010-0277/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cognition.2012.09.006

some of her other mental states (Was the agent able to see the victim? Was the action deliberate?). In addition, Greene and colleagues have argued that if and when a conflict arises between an intuitive/emotional and a more costly/non-emotional response to an action, the former will prevail, unless strong executive resources are available (Greene, 2009; Greene, Nystrom, Engell, Darley, & Cohen, 2004). We should therefore expect that under cognitive load, a relatively shallow non-mentalistic analysis of the immediate causes of the victim's suffering should prevail and that judges should be more severe in evaluating a case of accidental harm than if fuller cognitive resources were available for a complete appraisal of the situation.

Indirect evidence suggest that this prediction is plausible: Young and collaborators found that the presentation of accidental harm scenarios generates an increased activation in regions associated with cognitive conflict (Young et al., 2007). This reinforces the view that accidental harm is a type of conflict that requires available cognitive/executive resources for its resolution. Other empirical evidence come from the developmental literature: even though young toddlers and even infants are reliably sensitive to agents' goals, beliefs and intentions (Behne, Carpenter, Call, & Tomasello, 2005; Brandone & Wellman, 2009; Onishi & Baillargeon, 2005) and children's moral judgments have been shown to become sensitive to agents' intentions between the age of 3 and 5 (Nelson, 1980; Nelson-le Gall, 1985; Nobes, Panagiotaki, & Pawson, 2009; Shultz & Wright, 1986; Zelazo, Helwig, & Lau, 1996), children are also notorious for blaming accidental agents until 7–9 years of age (Hebble, 1971; Imamoğlu, 1975; Shultz & Wright, 1986; Piaget 1965/1932). This suggests that integrating intentions into moral judgments is a challenging task for children, especially when an agent's intention and her causal role conflict.

Even though it is plausible that cognitive load modulates the influence of intentional cues during moral evaluation, there is surprisingly little or no direct empirical demonstration that it does so in healthy adults. If such a modulation was documented, it would support the existence of a non-mentalistic cause-based heuristic in moral evaluation. Such a heuristic would take as input a causal description of a social interaction together with emotional cues, and output a negative evaluation of the agent who caused harm to a victim, irrespective of his or her intentions. Overriding such a heuristic for the purpose of evaluating an agent of accidental harm would require additional cognitive resources. To test this hypothesis, we designed two experiments. In Experiment 1, we presented non-verbal animated cartoons to two groups of adults, one of whom had to simply watch the cartoons, and the other of whom had, in addition, to perform a demanding verbal shadowing task (Forgeot d'Arc & Ramus, 2011; Hermer-Vazquez, Spelke, & Katsnelson, 1999; Newton & de Villiers, 2007). After seeing the cartoons, we asked them to evaluate two agents that only differed either in their causal contribution to the victim's suffering or in their intention to harm the victim. We expected intentions to prevail over causes, but only when cognitive resources were available. In Experiment 2, we tested the extent to which the dual task could also impair participants' ability to perceive or decode the agent's intention or causal role in these scenarios.

2. Experiment 1

We constructed three minimally different computer animated scenarios. In the *Coincidence* scenario, the agent is coincidentally present when the patient hurts himself. In the *Accident* scenario, the agent unwillingly harms the patient. In the *Aggression* scenario, the agent intentionally harms the patient. In all three scenarios, the victim suffers the same painful outcome, while the agent's movements are carefully matched. Healthy adults were distributed randomly into two groups. One group was required to compare the *accident* and the *coincidence* scenarios (the causal contrast). These scenarios only differed in whether the agent causes the victim's suffering, yielding a measure of the influence of the agent's causal role in moral/social evaluations. Participants in the second group were required to evaluate and compare the agent in the *accident* and the *aggression* scenarios (the intentional contrast) by answering a moral/social questionnaire, yielding a measure of the influence of intention ascription in moral/social evaluations. Half of the participants in each group had to perform a concurrent verbal shadowing task and half did not (No-load vs. With Load groups, respectively). None of these scenarios included any verbal content and all of the relevant causal and intentional variables had to be inferred from the movie. We expected that without cognitive load, adults would be more sensitive to the intentional than to the causal role of agents. Under cognitive load, however, they should display a pattern dictated by the cause-based heuristic, being thus more sensitive to the causal than the intentional contrast.

2.1. Materials and methods

2.1.1. Stimuli

Two versions of each of the three scenarios were constructed using Adobe Flash 8.0, one in which the agent is Mr. Green, the other in which the agent is Mr. Blue, yielding six animated clips, each lasting 10 s. In all clips, the agent (Mr. Blue or Mr. Green) is swinging near a road (on a swing or a rope), depending on the version. In the *aggression* clips, the agent faces the road, swings just once and stops. Then, he looks at the road as the victim (Mr. Red) is approaching and starts swinging again when Mr. Red stands right in front of him, intentionally hitting him. In the *accidental* clips, the agent is facing away from the road. He looks at the road while there is nobody (for the same duration as in the aggression clips), and starts swinging. While the agent is swinging, Mr. Red who is walking by is accidentally hit by the second swinging action. In the *coincidental* clips, the agent's movements are identical to those displayed in the accidental clips except that they are shifted in time (0.5 s), so that he stops swinging before the victim tumbles by himself (see Fig. 1). The stimuli were validated by conducting a pilot experiment (see Supplementary Section S1 for procedure and results).