Dispatch

Developmental Psychology: A Precursor of Moral Judgment in Human Infants?

P. Jacob[1] and E. Dupoux[2,3]

Human infants evaluate social interactions well before they can speak, and show a preference for characters that help others over characters that are not cooperative or are hindering.

The field of developmental psychology has advanced tremendously over the past thirty years, progress that is well illustrated by the recent paper of Hamlin *et al.* [1]. Pioneering work of Elisabeth Spelke, Susan Carey, Renée Baillargeon, Karen Wynn and many others has shown that, far from being immersed into William James' world of "booming and buzzing confusion", preverbal infants have highly structured expectations about the world: they parse the world into discrete and countable objects with properties like solidity and continuity through space and time [2,3]. Infants further distinguish between inanimate and animate objects: the former are governed by the law of Cartesian physics; the latter are self-propelled and infants take them to be rational agents of goal-directed behaviors [4,5]. These discoveries have been made possible by two major steps: first, the assumption that cognitive development is based not on general-purpose principles of associative learning, but rather on genetically determined, domain-specific acquisition systems [6–9]; and second, the emergence of astute experimental designs, capable of probing preverbal infants' behavioral reactions in response to their perception of simple *versus* complex, old *versus* new, or possible *versus* impossible, events — providing insight into their perception, memory and expectations [10].

So far, the social and moral world of preverbal infants has remained pretty much *terra incognita*. Past studies by Piaget [11], Kohlberg [12] and others have described human infants as being self-oriented or egocentric, or only responsive to adults' authority. But these studies used either informal and anecdotal observations or verbal reports, which are not readily usable before the age of three years old. Using two nonverbal experimental techniques, Hamlin *et al.* [1] have now shown that infants can evaluate a geometrical, cartoon-like agent involved in either helping or hindering another character who is trying to climb a hill (Figure 1). More specifically, a preference-choice technique shows that 10-month-old and even 6-month-old infants display a preference for the helping agent over the hindering one, and a violation of expectation paradigm shows that the 10-month-olds are more surprised to see the climber display a preference for the hinderer over the helper.

These result mesh well with the fast growing evidence that moral and social cognition is based on so-called 'core systems' — computationally specialized systems which process, in an automatic and unconscious fashion, evolutionarily relevant social and emotional information [13]. First, adults make very quick moral judgments, but, as Haidt [14] has shown, they often find themselves utterly dumbfounded when asked for explicit justifications. Secondly, as Blair [15] has argued, individuals who fail to empathize with the emotional distress of others, develop psychopathic behavior and are impaired in moral evaluation. Thirdly, as emphasized by Mikhail [16] and Hauser [17], moral judgments may require the representation of the intentional structure of social interactions, the computation of which is based in turn on deeply unconscious abstract principles. Finally, animal studies provide some evidence for the existence of complex social behaviors such as punishing cheaters [18] and comforting distressed conspecifics [19], which in humans may reflect moral judgments.

Does this mean that preverbal infants entertain moral thoughts? As Hamlin *et al.* [1] rightly put it, their experiment merely demonstrates that 6-month-olds display preferences for agents who help, rather than hinder, some unrelated third-party. Awareness of the work in

other domains of cognitive development, however, should make us suspicious of the claim that there is a single cognitive foundation for human moral cognition. Indeed, in the case of numerical cognition, there is not a single preverbal core system for numbers, but at least two: one system for tracking a small number of objects, and another system for evaluating large quantities or amounts of stuff [20]. Note that neither of these systems is genuinely numerical in the sense of number theory. Similarly, one could propose that morality rests not on one, but on several, core systems, none of which is intrinsically moral.

In the situation used by Hamlin *et al.* [1], there are at least two potentially confounded factors. The hindering agent is both frustrating the climber's intention and also potentially harming him by harshly pushing him on the slope. Similarly, if and when the helping agent promotes the climber's goal *after* the hindering agent has performed his negative act, then the question arises whether the helping agent might also provide comfort to the climber. In the abstract fo their paper, Hamlin *et al.* [1] tend to confuse two different social contrasts: helping *versus* hindering, and comforting *versus* harming. These two dimensions are dissociable: one can hinder the act of another agent so as to prevent him from harming himself. Conversely, one can help another agent perform a harmful act onto himself. In addition, these two dimensions may elicit different emotions: hindering an agent's act causes the agent's anger; harming a patient causes the patient's distress. If so, then there are reasons to think that these two dimensions are processed by two separate systems. Further research is needed to elucidate the number of separate social dimensions that are relevant for human infants.

Before closing, we would like to comment on the developmental difference found by Hamlin *et al.* [1] between the 6-month-old and the 10-month-old infants. They found that the 6-month-olds showed a preference for the helper over the hinderer, but were not more surprised to see the climber approach the hinderer as opposed to the helper. The 10-month-olds reacted in both tasks. The authors imply that there could be a developmental trend whereby infants would first use their own first-personal emotional responses in order to evaluate social interactions involving unrelated parties. Only later do they become able to represent the social evaluation of an agent by another character who was either helped or hindered by the agent's act. We suggest that the lack of emotional cues in the climber's responses to either the helper's positive act or the hinderer's negative act might explain the fact that 6-month-olds failed to display more surprise when they saw the climber join the hindering agent than the helping agent. At least, it is worth testing whether adding emotional cues on the part of the climber might enhance the surprise of 6-month-olds.

In brief, the findings by Hamlin *et al.* [1] raise several fascinating issues: is there a unique capacity for social evaluation or several? What is the link between the ability to evaluate helping vs. hindering agents and culturally acquired moral beliefs and norms regarding social cooperation? Are these systems partly learned on the basis of early social interactions? Or are they genetically pre-wired? Are non-human animals able to discriminate between helping and hindering agents too?

References
1.   Hamlin, J.K., Wynn, K., and Bloom, P. (2007). Social evaluation by preverbal infants. Nature *450*, 557-559
2.   Spelke, E.S. (1990) Principles of object perception. Cogn. Sci. *14*, 29-56
3.   Carey, S., and Xu, F. (2001). (2001) Infants knowledge of objects: Beyond object-files and object tracking. Cognition *80*, 179-213
4.   Premack, D. (1990). The infant's theory of self-propelled objects. Cognition *36*, 1-16
5.   Gergely, G., Nádasdy, Z., Csibra, G., and Biró, S. (1995). Taking the intentional stance at 12 months of age. Cognition *56*, 165-193

6.   Gallistel, C.R. (2000). The replacement of general-purpose learning models with adaptively specialized learning modules. In The Cognitive Neurosciences (2nd edn) (Gazzaniga, M.S., ed), 1179-1191, Cambridge, MA.

7.   Chomsky, N. (1975). Reflections on language. Pantheon Books

8.   Carey, S., and Spelke, L. (1994). Domain-specific knowledge and conceptual change. *In Mapping the Mind: Domain Specificity in Cognition and Culture (*Hirschfeld, A., and Gelman, S.A., eds), 169-200, Cambridge University Press

9.   Spelke, E.S., and Kinzler, K.D. (2007). Core knowledge. Dev. Sci. *10*, 89-96

10.   Mehler, J., and Dupoux, E. (1994). What infants know. Blackwell

11.   Piaget, J. (1932). Le jugement moral chez l'enfant. Presses Universitaires de France

12.   Kohlberg, L. (1981). Essays in Moral Development, vol. 1: the philosophy of moral development. Harper and Row

13.   Dupoux, E., and Jacob, P. (2007). Universal moral grammar: a critical appraisal. Trends Cogn. Sci. *11*, 373-378

14.   Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. Psychol. Rev. *108*, 814-834

15.   Blair, J.R. (1995). A cognitive developmental approach to morality - Investigating the psychopath. Cognition *57*, 1-29

16.   Mikhail, J. (2007). Universal moral grammar: theory, evidence and the future. Trends Cogn. Sci. *11*, 143-152

17.   Hauser, M.D. (2006). Moral Minds : How Nature Designed our Universal Sense of Right and Wrong. Ecco Press

18.   Hauser, M.D. (1992). Costs of deception: Cheaters are punished in rhesus monkeys (Macaca mulatta). Proc. Natl. Acad. Sci. USA *89*, 12137-12139

19.   de Waal, F.B.M., and Yoshihara, D. (1983) Reconciliation and redirected affection in rhesus monkeys. Behaviour *85*, 224-241

20.   Feigsenson, L., Dehaene, S., and Spelke, E. (2004) Core systems of numbers. Trends Cogn. Sci. *8*, 307-314

[1]Institut Jean Nicod, Ecole des Hautes Etudes en Sciences Sociales, Departement d'Etudes Cognitives-Ecole Normale Supérieure, Centre National de la Recherche Scientifique, 29 rue d'Ulm, Paris 75005, France. [2]Laboratoire de Sciences Cognitives et Psycholinguistique, Ecole des Hautes Etudes en Sciences Sociales, Departement d'Etudes Cognitives-Ecole Normale Supérieure, Centre National de la Recherche Scientifique, 29 rue d'Ulm, Paris 75005, France. [3]Maternité Port-Royal, Faculté de Médecine Port Royal-Cochin, APHP, Université René Descartes, 123 boulevard de Port-Royal, F75014 Paris, France. E-mail: emmanuel.dupoux@gmail.com

Figure 1. The nonverbal experimental technique used by Hamlin *et al.* [1]
(A) Social interaction events shown to infants. The climber (red character) attempts to climb the hill twice, each time falling back to the bottom of the hill. On the third attempt, the climber is either bumped up the hill by the helper (yellow character, left panel) or bumped down the hill by the hinderer (blue character, right panel). In the violation of expectation task (B), infants' looking times are measured for two events: The climber moves from the top of the hill to sit with the character on the right (left panel) or the left (right panel). In the choice paradigm, infants are presented with two toys, the helper and the hinderer, and are asked to choose one. (Reproduced with permission from [1].)