



Core Moral Concepts and the Sense of Fairness in Human Infants

Luca Surian¹ · Eugenio Parise² · Alessandra Geraci³

Accepted: 18 March 2025
© The Author(s) 2025

Abstract

We review recent experimental studies relevant to assess the proposal that human infants possess a sense of fairness that relies on sociomoral knowledge. We propose that this knowledge may include a core concept of justice with four foundational aspects: impartiality, agency, obligatoriness and conflicting claims. Infants' and toddlers' looking times, manual preferences and spontaneous actions provide some evidence for the first three features. Very early-emerging sociomoral evaluations and expectations about resource distributions show that infants process morally relevant information about distributors and recipients, suggesting that they are sensitive to the agency and impartiality constraints. Early evaluations appear to be linked to third-party expressions of praise or admonishment and to the deliverance of rewards and punishment, providing initial support for the obligatoriness constraint. More work is needed to investigate the sensitivity to conflicting claims, to assess the universality of early emerging evaluation skills and to show how core concepts relate to the development of explicit judgments and beliefs about duties and rights.

Keywords Fairness · Core Knowledge · Moral Judgment · Justice · Infancy · Cognitive Development

On a lucky day, one of us was taking his daughter to her kindergarten. Suddenly she shouted: “It is not fair!” (original: “*Non è giusto!*”, in her native Italian). “The teacher has favorites in class!”. The temptation to ask her a test question on ‘authority independence’ (Turiel, 1983) was irresistible: “Well, Sofia, the teacher is the

Eugenio Parise and Alessandra Geraci contributed equally to this work.

✉ Luca Surian
luca.surian@unitn.it

¹ Department of Psychology and Cognitive Science, University of Trento, Trento, Italy

² CIMeC - Center for Mind/Brain Sciences, University of Trento, Trento, Italy

³ Department of Educational Sciences, University of Catania, Catania, Italy

boss in class, isn't she? So, why is it not right for her to have favorites, even if she likes to do so?" She pondered for a moment and then replied, "Because the other children will feel bad!". While still unable to tie her shoes by herself, she could already articulate a protest against authorities' unfairness!

How do children get to this so fast? Do they acquire an understanding of fairness at around 3 years (Engelmann & Tomasello, 2019; Vaish & Tomasello, 2023) through collaborative interactions with other children and adults (e.g., Kohlberg, 1971; Piaget, 1932/1965; Smetana, 1983), cultural inculcation (Prinz, 2006), and other domain-general learning mechanisms (e.g., Dahl, 2019; Mammen & Paulus, 2023)? Or, rather, is the ontogenetic process bootstrapped by innate sociomoral skeletal knowledge, evolved tacit principles and domain-specific developmental mechanisms (e.g., Baillargeon et al., 2015; Hamlin, 2013a; Premack & Premack, 1997; Sheskin et al., 2014; Sperber & Baumard, 2012; Wynn & Bloom, 2014)? The main goal of this article is to review recent evidence that bears on these questions and allows an initial assessment of the claim that humans possess an early emerging sense of fairness that relies on a core concept of justice. This knowledge would play a key role in the development of explicit reasoning in the domain of distributive fairness (Deutsch, 1975), as well as in other domains of justice such as retributive (dealing with the delivery of deserved punishments and rewards; Dunlea & Heiphetz, 2020) and procedural fairness (dealing with how to take decisions in order to treat people fairly; Shaw & Olson, 2014).

Core Moral Concepts

Early emerging, possibly innate, core concepts and abstract principles help infants' to spontaneously acquire, quickly and spontaneously, sophisticated competencies about numbers, objects, places, agents, social beings and mental states (Baillargeon et al., 2016; Carey, 2009; Leslie, 1994; Spelke, 2022). Some authors have broadened this core knowledge view to account for the acquisition of a moral competence (e.g., Baillargeon et al., 2015; Dwyer et al., 2010; Hamlin et al., 2007; Ting et al., 2020), positing innate principles of harm avoidance, fairness, group loyalty and authority that constrain infants' reasoning and evaluation of agents and actions (Bian et al., 2018; Geraci & Surian, 2011; Haidt & Joseph, 2007; Margoni et al., 2018; Pellizzoni et al., 2010; Woo & Hamlin, 2023). Others have rejected this proposal, emphasizing the effects of early experiences and the role of learning mechanisms (Sommerville, 2023), or the viability of alternative explanations based on other core knowledge systems (Spelke, 2022, 2023; for replies see Hamlin, 2023; Revencu & Csibra, 2023). The implausibility of an evolutionary and modular view has also been pointed out. Carey (2009; p. 8) for example, states: "evolution did not provide us with an input analyzer that identifies electrons, *justice*, or 3,462,179" (emphasis added).

Here we will review recent developmental findings that reveal the early emergence of what people, possibly anywhere in the world, consider a central component of morality: the ability to tell just from unjust actions (Mikhail, 2022). We will cite most of the relevant works on infants (individuals aged 1 to 18 months) and toddlers

(19-to-30 months-olds), as well as few studies on older children and adults. Due to space limitations, we will not review the evidence on non-human animals (e.g., Brosnan & de Waal, 2003), and will only briefly refer to relevant cross-cultural works on older children and adults belonging to small-scale societies and cultures classified as non-WEIRD (i.e., white, educated, industrialized, rich, and democratic).

People typically reason and talk about fairness and justice in their attempts to solve conflictual situations and to evaluate many different things including laws, institutions, procedures, traditions, societies, actions and persons. Ancient Greek, Roman and medieval philosophers saw justice (*'justitia'*) as one of the most important virtues and, in line with this tradition, some psychologists consider it now a major strength of character (Peterson & Seligman, 2004). For Rawls (1958; 1971), justice is the first virtue of social institutions and fairness is its essential feature. Shifting the attention from virtues to actions, the ability to reason explicitly and autonomously about justice has been the main focus of classic theories of moral development (Damon, 1977; Kohlberg, 1971; Piaget, 1932/1965). Fairness and justice are terms used, sometimes interchangeably, with many different meanings. To avoid some confusions, it may be useful to start by saying that here we will focus on fairness and justice as properties attributed to actions and people, mainly in contexts involving distributions of resources (distributive justice), but also in contexts involving procedures, rewards and punishments (procedural and retributive justice).

Following Aristotle, many theories have emphasized that the common-sense conception of justice implies an idea of equality and fairness (Sen, 2009; Crisp, 2014). We may find a very good candidate for a definition of this concept in a sixth-century AD codification of Roman Law, the *Institutes of Justinian* (Miller, 2021) where justice is defined as 'the constant and perpetual will to render to each his due'. In this definition, we can identify four important features: 1) *Impartiality (fairness)*: justice is connected with the impartial, application of a rule or principle, it is understood as the opposite of arbitrariness; 2) *Agency*: justice requires that there is an agent that alters the states of other agents, evaluations of fairness are therefore linked to the understanding of agents and the representation of their rights and duties (a natural disaster can be judged as bad, not as unfair); 3) *Obligatoriness*: justice involves an obligation to each person, it is demanded, not begged; and 4) *Conflicting claims and interpersonal conflicts*: people reason about justice typically in contexts that involve conflicts and conflicting claims about how to divide and share utilities, such as resources, freedom or opportunities. Does the empirical evidence provided by recent infant studies meet a criterion based on these features?

Impartiality

Numerous recent studies investigated infants' sense of fairness by presenting them with scenarios involving actions that violate or conform to *distributive justice* principles based on equality (egalitarian rules), or merit and deservingness (equity rules). According to Engelmann and Tomasello (2019), unfair actions may acquire their high negative value not out of a concern for the material benefits, but because they convey that someone is not getting the deserved respect

(in the sense of what Darwall (1977; p. 38) called ‘recognition respect’: “giving appropriate consideration or recognition to some feature of its object in deliberating about what to do”). Infants were shown events involving simple resource distributions (for example, a cartoon animal distributes colored disks or strawberries to two other agents) and their looking times, helping preferences or manual choices have been recorded. The recipients were cartoon animals (e.g., Geraci & Surian, 2011); human adults (e.g., Schmidt & Sommerville, 2011), puppets (e.g., Bian et al., 2018; Sloane et al., 2012) or self-moving geometrical shapes with eyes (e.g., Meristo & Surian, 2013). The Violation of Expectation (VoE) paradigm provided the rationale for most of these experiments: longer looks at events that violate an expectation or a principle, compared to events that conform to it, suggest that infants generated expectations or inferences based on such a principle (for a conceptual overview see Margoni et al., 2024). In order to interpret the observed patterns of looking times as resulting from an expectation violation and to rule out alternative, more conservative explanations based on affiliative biases or low-level factors such as familiarity and perceptual symmetry, VoE studies typically include various control conditions (for a cogent discussion, see Hamlin, 2023; see also Stahl & Feigenson, 2019), or rely on convergent neuroimaging evidence (e.g., Forgács et al., 2020; Hudac & Sommerville, 2020; Southgate & Vernetti, 2014).

Infants display a variety of reactions to events involving resource distributions. All things being equal, infants and toddlers look longer at unequal than at equal distributions, suggesting that they, like preschoolers (Fehr et al., 2008; Elenbaas, 2019), expect agents to be impartial and follow an egalitarian rule to distribute windfall resources (e.g., Buyukozur Dawkins et al., 2019; Geraci & Surian, 2023a; Meristo et al., 2016; Schmidt & Sommerville, 2011; Sloane et al., 2012). Moreover, their manual choices reveal a reliable preference for fair over unfair agents (e.g., Geraci & Surian, 2011; Lucca et al., 2018), suggesting that they tag fair agents more positively than unfair ones (see also Burns & Sommerville, 2014). Consistent with this preference, toddlers also selectively help (Surian & Franchin, 2017a), share their own goods (Nava et al., 2019), accept a toy from (Burns & Sommerville, 2014) and reward fair agents (Ziv et al., 2021) rather than unfair ones.

Some studies have included control conditions in which resources are revealed rather than distributed (e.g., Meristo et al., 2016; Sloane et al., 2012) and they have consistently found that infants do not look longer at unequal allocations, they look longer only at unequal distributions, suggesting that they are not expecting that similar individuals have similar amounts of resources, rather they generate expectations about the distributive actions.

Toddlers can also rely on an equity principle: their looking times reveal that they expect resources to be distributed according to recipients’ relative merit and deservingness (Sloane et al., 2012; Surian & Franchin, 2017b; Wang & Henderson, 2018). Moreover, their expectations are affected by information about the recipients’ social dominance and in-group membership. Toddlers expect a distributor to display ingroup favoritism when resources are scarce and an egalitarian sharing is not possible (Bian et al., 2018) and they look longer when a distributor favors a subordinate individual instead of a dominant one (Enright et al., 2017). This suggests that infants

expected the distributor to favor dominant recipients, but it is also possible that they were focusing on what the dominant character might do after being denied a larger share of the resources.

Children and adults care not only about distributive justice, but also about *procedural justice*: they are concerned both about the final outcome of a distribution and about how a certain outcome was achieved (e.g., Grocke et al., 2015; Jacobs et al., 2022; Shaw & Olson, 2014; Stowe et al., 2022). At present, only one study investigated whether also toddlers are sensitive to violations of procedural fairness (Surian & Margoni, 2020). This study, however, did not present toddlers with distributive actions but with scenarios in which two puppets asked for help in opening a box to a fair or to an unfair helper. Following the requests for help, the fair helper provided help to both puppets after the same delay. By contrast, the unfair helper provided help immediately to one puppet, but left the other puppet waiting (asking three or five times) before receiving the requested help. Children looked longer at the procedurally unfair event, suggesting that they expected an equal solicitude and detected a violation of fairness.

Infants' expectations of agents' fairness are affected by observing prior violations in different moral domains. After seeing an agent performing a helping action, 10-month-olds expected her to perform a fair distribution, but when the agent first performed a hindering action, they no longer expected her to distribute fairly (Surian et al., 2018). These results suggest that infants are able to form representations of agents' moral character (Sperber & Baumard, 2012). Convergent results were found by Ting and Baillargeon (2021): at 25 months, fairness expectations are affected by observing both previous hindering actions, as in Surian et al., (2018), and actions that violate the in-group loyalty principle. Moreover, the link between the domain of harm and the domain of distributive unfairness appears to be bidirectional. Gill and Sommerville (2023) found that witnessing an agent's distributive behaviors affected infants' expectations of whether that agent would later help or hinder another agent. Overall, several findings from different labs indicate that the impartiality constraint is taken into account by infants and toddlers.

Agency

The evaluation of fairness and justice should be constrained only to contexts and events in which agents modify the states of other agents. Some infant studies included control conditions in which recipients were replaced by inert objects or by inanimate puppets that do not move or talk. These conditions provide good controls as the inanimate puppets were perceptually identical to the self-moving puppets used in the experimental conditions. Infants in the first and second year saw an agent moving resources from one place to another place – like in the distributive events –, but those spatial changes did not modify the state of other agents (e.g., Geraci & Surian, 2011; Meristo et al., 2016; Sloane et al., 2012; Strid & Meristo, 2020). Consistently with the *agency* constraint, in these scenarios infants and toddlers did not show a preference for agents that performed symmetrical motions.

Infants individuate agents by attending to simple perceptual cues, such as autonomous motion (Decarli et al., 2020; Surian & Caldi, 2010). The evaluation of fairness requires the individuation of agents and such individuations and categorizations processes may exploit the representational resources provided by infants' actional understanding (Leslie, 1994) and the agents core system of knowledge (Spelke, 2022). But studies on fairness evaluation suggest that infants represent agents as also endowed with sociomoral properties (Ting et al., 2020). Therefore, the representations of agents that infants generate during their evaluations appear to be richer than those provided by the core systems proposed by Spelke (2022).

The agency constraint requires that fairness evaluations should be linked to the individuation and understanding of agents. By studying how infants' evaluations of distributions are intertwined with their understanding of agents one can gain further assessments of the agency constraint and shed some light on the mechanisms underpinning infants' responses. One way to reason about agents is to attribute to them goals and preferences and take into account relevant physical constraints (Gergely & Csibra, 2003; Leslie, 1994; Woodward, 1998). Preverbal infants understand actions goal-directedness and react less negatively towards an adult that is unable to give them a toy than towards one that is teasingly unwilling to do so (Behne et al., 2005; Geraci et al., 2024; Luo, 2011; Sommerville et al., 2005). At 15 months they ascribe the goal of giving an object by taking into account cues about object possession (Tatone et al., 2019). Crucially, infants' ability to attribute goals and preferences predicts, 5 years later, their intention-based moral judgments: children who, at 7 months, passed a well-known goal-attribution task (Woodward, 1998), performed better, at the age of 5 years, in a task involving intention-based moral judgments (Sodian et al., 2016).

Generating moral judgements by attending to agents' intentions is a universal aspect of moral competence (Barrett & Saxe, 2021; Mikhail, 2022) that emerges very early in development (see Hamlin, 2013b, for findings in the help/harm domain). Preverbal infants attend not only to distribution outcomes, but also to distributors' intentions. Strid and Meristo (2020) showed 10-month-olds simple scenarios involving one distributor and two recipients. After giving one strawberry to one of the recipients, the distributor failed to deliver the second strawberry either to the same or to the other recipient. In the test events, a third party approached one of the two distributors, and infants looked longer at the approach-unfair event, suggesting that they considered the distributors' intentions and expected the third party to approach the distributor who tried to act fairly. Two further studies presented failed attempts to distribute windfall resources fairly and unfairly and found that young infants, aged 9 (Geraci et al., 2022) and 4 months (Geraci & Surian, 2023a) preferentially chose and looked at the fair distributor. This very early emergence satisfies the age requirement for core knowledge (Spelke, 2022).

Many studies found evidence showing that infants can go beyond the attribution of goals (or intentions) when reasoning about agents, they also take a mentalistic stance by attributing epistemic mental states (e.g., Choi & Luo, 2015; Onishi & Baillargeon, 2005; Surian et al., 2007). Mentalistic reasoning typically involves the representation of agents' true and false beliefs (i.e., metarepresentations) to predict or make sense of their actions. Most experimental studies revealed

metarepresentational skills in the second year (for reviews see Baillargeon et al., 2015; Scott & Baillargeon, 2017; Kamps et al., 2020; Sodian et al., 2020), but some studies reported positive evidence of mentalistic understanding also in pre-verbal infants (Choi et al., 2022; Kamps et al., 2015; Kovács et al., 2010; Southgate & Vernetti, 2014). Sloane et al. (2012) and Meristo and Surian (2013) found evidence suggesting that infants can reason in a mentalistic way when they process scenarios involving resources distributions. Sloane et al. (2012) presented a task designed to assess toddlers' merit-based evaluations. Sloane et al. (2012) found that toddlers generally expect compensations to be commensurate with efforts. However, participants did not show evidence of having these expectations when the distributor ignored relevant information. In that study, toddlers first watched a distributor assigning a task to two individuals (e.g., tidying up the room). Then she performed an equal distribution towards the same individuals, while ignoring that only one of them had obeyed and completed the task. In Meristo and Surian (2013) study, ten-month-old infants were shown an agent rewarding a fair or an unfair distributor. When it was clear that the rewarding agent had watched the distributive actions performed by the distributor, infants were surprised to see that she rewarded the unfair distributor. By contrast, when it was clear that the rewarding agent could not have seen the distributive actions, infants had no specific expectation about which distributor would receive the reward. These findings suggest that infants and toddlers took into account distributors informational states to generate relevant behavioral expectations (for related results showing infants' mentalistic reasoning in the evaluations of hinderers and helpers, see Choi & Luo, 2015; Hamlin, 2013b; Woo et al., 2017; Woo & Spelke, 2022).

In sum, several findings indicate that infants take into account the agency constraint when they evaluate distributive actions. Moreover, the links between infants' early sociomoral evaluations and their ability to reason about agents' intentions and epistemic mental states provide some support for models that emphasize developmental continuity rather than radical conceptual changes (Margoni & Surian, 2017).

Obligatoriness

The feeling of obligation may have its roots in reciprocity mechanisms that allowed the evolution of cooperative hunting in our Pleistocene ancestors and provided the foundations for the ability to learn social norms quickly and spontaneously (Tomasello et al., 2012). Expectations of reciprocity and fairness emerge early in human infants (Jin et al., 2024) and rely on principles that are quite different from those that operate in other domains. While abstract principles in the core domains of objects, places and numbers lead to expectations about what is likely to occur in a given context, sociomoral principles mandating fairness and prohibiting harm generate prescriptions about what ought to occur and how to evaluate actions as good and right, or bad and wrong. The *obligatoriness* constraint is a crucial aspect of a core concept of justice, but it is a very difficult one to assess in infants and toddlers. For this reason, it is a source of long-lasting controversies. Many studies showed that

infants prefer prosocial to antisocial individuals and expect equal or merit-based distributions, but we need to distinguish preferences and expectations from feelings of obligation. An evaluation of fairness and justice brings with it a sense of obligation that preferences or expectations do not necessarily include.

Infants selectively associate fair or unfair actions with expressions of praise or admonishment (DesChamps et al., 2016). Infants at 15 months watched a fair and an unfair distributor side by side and, while they were hearing expressions of praise, they looked longer at the unfair agent than at the fair one, showing a reliable mismatching effect. By contrast, when they were hearing admonishments they showed no reliable visual preference for one of the two agents. At 13 months, infants showed a reliable preference for the unfair agent while they were listening to admonishments, showing a matching rather than a mismatching effect. Infants also link fair and unfair actions with the deliverance of rewards (Meristo & Surian, 2013) and punishment (Geraci & Surian, 2023b). In Meristo and Surian (2013) study, 10-month-olds were surprised to see that a third-party chose to reward an unfair distributor (by giving a strawberry). Likewise, Geraci and Surian (2023b) found that 9-month-olds infants, in test trials that showed simultaneously a fair and an unfair distributor, were surprised when a third-party chose to reward the unfair agent. Infants were also surprised when they watched the third-party chose to deliver a physical punishment to the fair rather than the unfair distributor by hitting him with a stick. These findings provide further evidence revealing the sociomoral nature of infants' evaluations and suggesting an early-emerging understanding of retributive actions. Further evidence in line with this conclusion was reported by Ziv et al. (2021) who found a spontaneous tendency to deliver rewards to fair rather than to unfair agents by selectively touching the area of a screen that activated a recording of vocal expressions of praising. Consistently, Stavans and Baillargeon (2019) found that 17-month-olds expect a leader to intervene and rectify the situation when an ingroup member takes more than its share of a windfall resource, suggesting a sense that a wrong has been committed that needs to be corrected. Moreover, they hold no such expectation for a non-leader, who might be scared of reprisals from the wrongdoer.

These reactions suggest that infants and toddlers hold expectations about third parties' retributive behaviors enacted towards fair and unfair distributors and they can even act to deliver such retributions. Is this showing that infants link fairness to obligatoriness? Only weakly, since punishments and rewards do not necessarily derive from moral disapprobation or approval, they can also be delivered, more simply, out of dislike or a wish for behavioral control.

For some authors, a feeling of obligation can only be revealed by explicit verbal judgments, justifications, and protests that display a normative vocabulary (Dahl, 2018; Paulus, 2020). Notice, however, that even if you subscribe to this view, you still need to clarify what young children mean when they use normative and evaluative terms in their utterances (Franchin et al., 2019), and explain how they succeed in acquiring the semantics of those abstract words. One possibility is that the early emergence of some core sociomoral concepts allows toddlers to fast map the meaning of their first normative and evaluative words. Further evidence on the early emergence of a sense of obligation may be obtained, in future studies, by recording

how infants react to people performing actions that we see as obligatory or as super-erogatory (McManus et al., 2020; Ting, 2020). Overall, the existing evidence is consistent with an early sensitivity to obligatoriness, but the proofs are not conclusive and need to be strengthened in future works.

Conflicting Claims and Interpersonal Conflicts

Justice and fairness are usually invoked when conflicting claims and interpersonal conflicts arise and when people try answer questions like ‘How one ought to divide these resources given the existing conflicting claims?’. By contrast, when agents’ goals do not compete, but rather converge, efficiency considerations may take central stage (‘What is the best way to reach our shared goals?’). Also, David Hume claimed that in a hypothetical state of abundance where any conflict would turn out useless “the jealous virtue of justice would never once have been dreamed of” (Hume, 1975, p. 184; cit. in Miller, 2021; see also Rawls, 1971). So, real or hypothetical interpersonal conflicts about resource distributions, welfare and rights, are probably the typical contexts in which the concept of justice is invoked. Is there any evidence, in the infant literature, for a link between conflicts and the activation of the concept of justice? There is abundant evidence that infants attribute goals, intentions and other mental states to agents, as we discussed above, and there is also some evidence that they, by 10–13 months, can represent interpersonal conflicts (e.g., Mascaro & Csibra, 2012; Thomsen et al., 2011) but, at present, we have no proof that their ability to detect conflicts is linked to the evaluation of distributive fairness.

It is plausible that, when infants are faced with the simple scenarios used to assess their sense of fairness, they represent the recipients as having an interest in receiving the available resources and as having conflicting claims, goals or desires about the final outcome of the distribution. However, direct evidence for this has not been reported yet. It would be therefore of great value to study whether the presence of a salient conflict between agents and claims primes infants’ evaluations of fairness. Another way to investigate this question would be to present infants with scenarios in which it is clear that there are no conflicting claims and test whether this would affect their evaluations and expectations.

The fact that researchers have, so far, paid less attention to the conflicting claims component than to the other three components may not be without a reason. Conflicting claims do not appear to be, like the other three aspects, a necessary component of JUSTICE, but rather a contextual feature affecting its instantiation. While the need to solve social conflicts is likely to have played, in our ancient past, a crucial role in the evolution of a sense of fairness (Tomasello et al., 2012), we do not know whether the perception of conflicting claims, or an interpersonal conflict, facilitates the activation of the concept of justice. To test this possibility, future studies could build on what we know about infants’ understanding of social power and conflict (Margoni et al., 2018; Pun et al., 2016).

A Comparison with Related Theoretical Proposals

Our proposal about infants' sense of fairness and the core concept of justice dovetails with some previous proposals put forward by leading authors in this research field. Baillargeon, Bian, Buyukozer Dawkins, Ting, and their colleagues claimed that early moral cognition includes four moral principles (fairness, harm avoidance, in-group support, and authority) and four dedicated moral concepts (obligations, status, circle, and character; Baillargeon et al., 2015; Buyukozer Dawkins et al., 2019; Ting et al., 2020). These principles are similar to those posited in adult studies by the Moral Foundations Theory (Graham et al., 2013; Haidt, & Joseph, 2007) and they do not determine who is virtuous, but what actions are obligatory, forbidden or permissible, thus generating expectations about what an individual will (and should) do in certain situations. Infants possess a concept of moral obligation and rely on an "equity-based principle of fairness: all other things being equal, individuals are expected to give others their just deserts—that is to treat them as they deserve to be treated in the situation at hand" (Ting et al., 2020; p. 44). These principles apply to a variety of situations, including those that involve the division of windfall resources, the dispensation of rewards to individuals commensurate to their relative effort, and the delivery of punishments to disobedient agents. Also, "infants conceptualize the various principle-based expectations as *obligations* that specify how individuals should act toward others (e.g., individuals should act fairly)" (Ting et al., 2020; p.49, emphasis in the original). All this is coherent with the proposal we put forward in the present paper. So, what do we gain by positing also a core concept of justice? The principle of fairness, proposed by Baillargeon and her colleagues, prescribes that individuals should act fairly. Needless to say, here 'should' means 'ought to', so such a principle requires the concepts of INDIVIDUAL, OUGHT, and ACTING FAIRLY. Ting et al. (2020) discuss in depth the first two (INDIVIDUAL, as an agent endowed of moral status, and OBLIGATION), but not the concept of JUSTICE as a property of actions and agents. This concept is one of the necessary building blocks of the principle of fairness. Thus, our proposal is not meant to replace, but rather to integrate their broader model.

Overall, our proposal is also coherent with many works by Hamlin, Bloom, Tan, Woo and Wynn (Bloom, 2014; Hamlin, 2013a, b; Hamlin et al., 2007; Wynn & Bloom, 2014; Tan & Hamlin, 2022; Woo et al., 2022; Tan et al., 2018) where we read that "humans' core system of knowledge include primitive abilities to understand and evaluate morally relevant actions (...). We term these abilities a 'moral core'." (Woo et al., 2022; p. 41). While their experimental work has been focused, primarily, on infants' evaluation of actions that help or hinder other agents' attempts to fulfill their goals, our work has been focused mainly on the domain of distributive fairness and it expands, in a consistent way, their proposals. However, in reviewing the available evidence on infants' evaluation in different moral domains, Woo and Hamlin (2022) reached the conclusion that while there are many findings showing a preference for helping agents (over hindering ones) in the first year of life (e.g., Hamlin et al., 2010), a preference for fair agents has been found in the second, but not in the first year (Burns & Sommerville,

2014; Geraci & Surian, 2011; Lucca et al., 2018), and this suggest a different developmental onset for the principles prohibiting harm and prescribing fairness. Since recent works found expectations of fairness at 4 months (Buyukozer Dawkins et al., 2019) and a preference for fair individuals in infants aged 4 and 9 months (Geraci & Surian, 2023a, 2023b; Geraci et al., 2022), we believe that the available evidence suggests that both the harm-avoidance and the fairness principles emerge in the first year, at approximately the same time. Moreover, we suspect that these two domains may be more closely linked than it has been assumed by the Moral Foundations Theory (Graham et al., 2018). It is possible that the delivery of an unfair distribution is always perceived, implicitly and automatically, as involving harm, similarly to what happens when people read scenarios involving impure actions, even if such actions are described and judged explicitly as totally safe and harmless (e.g., Gray et al., 2014; Osler et al., 2024). In this light, it would also be useful to investigate, in future studies, the links between infants' sense of fairness and the early emergence of concerns for others (Davidov et al., 2013).

Sommerville, Lucca, Burns, Enright, Ziv and their colleagues reported many seminal findings that are consistent with our view, but there are also some noteworthy differences between their theoretical perspective and ours. They reported infants' expectation of fairness (Schmidt & Sommerville, 2011; Sommerville et al., 2013; Enright et al., 2017) as well as infants' preference for fair agents (Burns & Sommerville, 2014; Lucca et al., 2018; see Sommerville, 2018, 2023, for reviews), but they also found some developmental changes that, in their view, support a constructivist perspective (Ziv & Sommerville, 2017). We instead think that those findings do not support selectively neither a constructivist nor a nativist view, as they are mute about the domain generality or specificity of the mechanisms underlying the early acquisition of sociomoral preferences. By contrast, the very early emergence of sociomoral evaluations, expectations and preferences do provide some evidence for the nativist views and they are at odds with constructivist theories, particularly with those that emphasize the role of peer interactions (Killen, 2018).

The domain specificity of the acquisition mechanism is a central issue in any attempt to adjudicate between nativist and constructivist models of cognitive and moral development (Laurence & Margolis, 2024). A sophisticated attempt to take seriously this issue can be found in the authors that endorse the linguistic analogy (Rawls, 1971) and posit the existence of a Universal Moral Grammar (Mikhail, 2007). Like the Universal Grammar, proposed to account for native language acquisition, the Universal Moral Grammar would provide the representational resources needed to parse morally relevant actions and contexts, allowing children to rapidly and spontaneously acquire the morality that characterizes the culture in which they are immersed. This view focused on representations needed for deontic reasoning and the judgment of intentional and accidental harming actions as prohibited or permissible, as a function of their outcomes and intentions (e.g., Dwyer et al., 2010; Levine et al., 2018). While our proposal is consistent with this account, more theoretical work and a great deal of methodological ingeniousness are necessary to show how this approach can be fruitfully extended to account for infants' sense of fairness.

Future Directions

Simple egalitarian expectations appear very early, in infants as young as 4 months, and merit-based evaluations emerge in toddlers. This very early emergence and fast development is more consistent with an initial state that is endowed with core knowledge and domain-specific developmental mechanisms than with views that deny any knowledge of abstract principles and posit only slow, domain-general associative learning. The evidence we reviewed here provides some initial support for an early emerging core concept of justice. It does not rule out, however, a role for constructivist developmental processes that may be responsible for important conceptual changes occurring later, during childhood and adolescence (Killen et al., 2022). Also, we cautiously say that this evidence should be seen as providing ‘initial support’ for a core concept of justice since it is neither unequivocal nor conclusive. More work is needed to consolidate these findings via multi-lab replications (e.g., Lucca et al., 2021), to deepen our understanding of the information processing mechanisms underpinning infants’ evaluations and expectations (Stahl & Feigenson, 2019), and to further assess the role of environmental input and statistical learning (Ruffman et al., 2012). Longitudinal studies could fruitfully clarify the present findings by testing whether infants’ reactions to unfair distributions reliably predict the judgments that children will express verbally some years later (see Tan et al., 2018 for related results).

Relevant cross-cultural infant studies are, unfortunately, still rare. Only one published study assessed toddlers’ sense of fairness in non-WEIRD populations from small-scale societies and it reported remarkable cross-cultural differences. Meristo and Zeidler (2022) compared a group of 12 to 20-month-olds from Sweden with infants from Samburu and Kikuyo populations, two Kenyan ethnic groups. Swedish infants looked longer at unequal distributions, whereas Kikuyo infants looked equally at the two types of distributive outcomes and Samburu infants looked longer at equal distributions, suggesting an early effect of culture-specific experience. Further cross-cultural infant studies could be inspired by previous research on adults and children. Concerns for fairness and egalitarian principles in distributions appear to be a universal aspect of human moral competence (Boehm, 2012; Curry et al., 2019). However, a rich literature on adults’ economic reasoning and decision making has also revealed noteworthy cross-cultural variations on how people judge distributions and how they decide to share windfall resources (e.g., Henrich et al., 2005). Participants from WEIRD and from many small-scale societies may opt for an egalitarian rule when they participate in a sharing task, or when they are asked to judge the fairness of a distribution, but, as a second choice, children from the US favor merit-based rules while children from other cultures may favor need-based rules (e.g., Carson & Banuazizi, 2008; see also Huppert et al., 2019). A universal, early emerging principle of justice may mandate that one should treat like cases alike and different cases differently. If this principle is so formulated, it lacks determinant content, but it helps to order our experience, allowing the construction of a “theory of justice of which features of similarities and differences actually matter” (Shweder, 2004, p. 91).

Cross-cultural differences in behavioral choices are predicted by social factors such as the level of market integration (i.e., how often an individual is involved in buying or selling goods; Henrich et al., 2005). Young children from a great variety of cultures tend to behave more selfishly than adults in their sharing decisions, by keeping most of the resources for themselves (Cowell et al., 2017; Rochat et al., 2009), but by the age of 5 years they possess a sense of ownership that leads them to spontaneously respect others' properties (Kanngiesser et al., 2019) and by mid-childhood they follow egalitarian rules and spontaneously choose to punish agents that behave selfishly in resource distributions (House et al., 2013, 2020). While we know that infants' expectations of fairness are linked to their sharing decisions (Schmidt & Sommerville, 2011), and early sharing experiences predict infants' later reactions to fair and unfair distributions (Ziv & Sommerville, 2017), there is much to study, both longitudinally and cross-culturally, to better describe and understand these links. No previous study has assessed whether infants can follow need-based principles to evaluate resource distributions and whether they can reason about procedural fairness in distributive contexts (Shaw & Olson, 2014).

A central and early emerging aspect of moral competence is the ability to tell moral norms from conventional norms, assuming that only moral norms are authority and context dependent (Turiel, 1983). Beliefs in the context and authority independence of moral norms are found in WEIRD as well as in small scale societies (Piazza & Sousa, 2016), although the expression of these beliefs by adults' is affected by pragmatic factors (Margoni & Surian, 2021). Even toddlers react differently to moral and conventional transgressions (Smetana, 1984). Do they also see distributive and procedural unfairness as violations of a moral norm? The evidence reviewed here appear to support a positive answer, but recent results on preschoolers suggest they do not (Yucel et al., 2022) and therefore more direct tests are necessary.

We still know very little about how infants face situations that may involve conflicts between different principles, for example a conflict between fairness and respect for authorities or group loyalty (Bian et al., 2018; Ting et al., 2020; Woo & Hamlin, 2023), or a conflict between merit-based and need-based distributive rules. The evidence we reviewed above shows that infants and toddlers evaluate the fairness of many types of agents, including non-human agents. We need further investigations of how fairness and justice concerns are constrained by information about ingroup/outgroup membership, moral circle membership and agents' hierarchical status.

Conclusions

Humans' success in adapting to the most diverse environments relies on the capacity to coordinate their actions with those of many other unrelated individuals and engage in sophisticated cooperative activities. We are 'super-cooperators' because, among other things, we can rely on abstract concepts that allow us to reason about agents and evaluate their actions, including their fairness (Baumard et al., 2013). It is plausible that the acquisition of these universal skills is underpinned by some

proximate mechanisms and some universal norms that “evolved because they constitute the most effective means of upholding mutually beneficiary cooperative relations” (Krebs, 2008, p. 168). By emphasizing the role of these capacities in allowing complex cooperation, these views help to identify moral universals (Blake et al., 2015; Curry et al., 2019) and to explain the developmental time course of fairness concerns and sharing preferences (Sheskin et al., 2014).

Consistently with this evolutionary view, a growing body of evidence reveals the early emergence of a sense of fairness. Infants’ and toddlers’ expectations and preferences suggest that their evaluations of fairness manifest most of the signature features that are also found in children’s and adults’ moral evaluations. These evaluations are restricted to agents, they concern violations of both distributive and procedural norms, they rely on agents’ intentions and other mental states, they are linked to expectations of third-party punishment and rewards, and they even motivate rewarding actions. Moreover, the involvement of theory of mind undermines the viability of alternative accounts that tried to explain looking times results as solely due to the effect of low-level perceptual factors.

In the past decade, many studies have consistently provided support for the claim that infants’ and toddlers’ sociomoral competence is much richer than previously assumed. Here we proposed that it may even include a rudimentary sense of fairness that relies on a tacit concept of justice. These proposals posit the ‘developmental roots’ or the ‘seeds’ of fairness, but they are not meant to revive preformistic views, deny that developmental changes occur in childhood and adolescence, or suggest that the development of sociomoral skills is merely the unfolding of a genetic blueprint. By clarifying and extending these results in longitudinal and cross-cultural studies, future research will have the chance to broaden our understanding of how humans acquire a mature moral competence and develop their full-fledged ability to reflect on persons’ duties and rights.

Acknowledgements We are grateful to Francesco Margoni, Marek Meristo, Priya Silverstein and two anonymous reviewers for their comments and suggestions on earlier drafts of this work.

Author Contribution L.S. had the idea for the article, performed the literature search, and wrote the first draft of the article. All authors comments on previous versions and critically revised the work.

Funding Open access funding provided by Università degli Studi di Trento within the CRUI-CARE Agreement. This work was supported by the Ministero dell’Università e della Ricerca (PRIN-PNRR grant no. P2022M9EKK to L.S.).

Data Availability We do not analyze or generate any datasets, because our work proceeds as a review of the current literature, but does not provide a meta-analysis of the results reported in previous articles. One can obtain the details on these results and, in some cases, the access to the original data sets, from the references below.

Declarations

Conflict of Interests We declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative

Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Baillargeon, R., Scott, R. M., & Bian, L. (2016). Psychological reasoning in infancy. *Annual Review of Psychology*, 67, 159–186.
- Baillargeon, R., Scott, R. M., He, Z., Sloane, S., Setoh, P., Jin, K., Wu, D., & Bian, L. (2015). Psychological and sociomoral reasoning in infancy. In M. Mikulincer, P. R. Shaver (Eds.), E. Borgida, & J. A. Bargh (Assoc. Eds.), *APA handbook of personality and social psychology: Vol. 1. Attitudes and social cognition* (pp. 79–150). Washington, DC: American Psychological Association.
- Barrett, H. C., & Saxe, R. R. (2021). Are some cultures more mind-minded in their moral judgements than others? *Philosophical Transactions of the Royal Society B*, 376(1838), 20200288.
- Baumard, N., André, J. B., & Sperber, D. (2013). A mutualistic approach to morality: The evolution of fairness by partner choice. *Behavioral and Brain Sciences*, 36(1), 59–78.
- Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling versus unable: Infants' understanding of intentional action. *Developmental Psychology*, 41(2), 328–337. <https://doi.org/10.1037/0012-1649.41.2.328>
- Bian, L., Sloane, S., & Baillargeon, R. (2018). Infants expect ingroup support to override fairness when resources are limited. *Proceedings of the National Academy of Sciences*, 115(11), 2705–2710. <https://doi.org/10.1073/pnas.1719445115>
- Blake, P. R., McAuliffe, K., Corbit, J., Callaghan, T. C., Barry, O., Bowie, A., ... & Warneken, F. (2015). The ontogeny of fairness in seven societies. *Nature*, 528(7581), 258–261.
- Bloom, P. (2014). *Just babies: The origins of good and evil*. Crown.
- Boehm, C. (2012). *Moral origins. The evolution of virtue, altruism and shame*. Basic Books.
- Brosnan, S. F., & de Waal, F. B. M. (2003). Monkeys reject unequal pay. *Nature*, 425, 297–299.
- Burns, M. P., Sommerville, J. A. (2014). "I pick you": the impact of fairness and race on infants' selection of social partners. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.00093>
- Buyukozur Dawkins, M., Sloane, S., & Baillargeon, R. (2019). Do infants in the first year of life expect equal resource allocations? *Frontiers in Psychology*, 10, 116. <https://doi.org/10.3389/fpsyg.2019.00116>
- Carey, S. (2009). *The origin of concepts*. Oxford University Press.
- Carson, A. S., & Banuazizi, A. (2008). "That's not fair": Similarities and differences in distributive justice reasoning between American and Filipino children. *Journal of Cross-Cultural Psychology*, 39(4), 493–514.
- Choi, Y. J., & Luo, Y. (2015). 13-month-olds' understanding of social interactions. *Psychological Science*, 26(3), 274–283. <https://doi.org/10.1177/0956797614562452>
- Choi, Y., Luo, Y., & Baillargeon, R. (2022). Can 5-month-old infants consider the perspective of a novel eyeless agent? New evidence for early mentalistic reasoning. *Child Development*, 93(2), 571–581.
- Cowell, J. M., Lee, K., Malcolm-Smith, S., Selcuk, B., Zhou, X., & Decety, J. (2017). The development of generosity and moral cognition across five cultures. *Developmental Science*, 20, e12403. <https://doi.org/10.1111/desc.12403>
- Crisp, R. (Ed.). (2014). *Aristotle: Nicomachean Ethics*. Cambridge University Press.
- Curry, O. S., Mullins, D. A., & Whitehouse, H. (2019). Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. *Current Anthropology*, 60(1), 47–69.
- Dahl, A. (2018). New beginnings: An interactionist and constructivist approach to early moral development. *Human Development*, 61(4–5), 232–247. <https://doi.org/10.1159/000492801>

- Dahl, A. (2019). The science of early moral development: On defining, constructing, and studying morality from birth. *Advances in Child Development and Behavior*, 56, 1–35. <https://doi.org/10.1016/bs.acdb.2018.11.001>
- Damon, W. (1977). *The social world of the child*. Jossey-Bass.
- Darwall, S. L. (1977). Two kinds of respect. *Ethics*, 88, 36–49.
- Davidov, M., Zahn-Waxler, C., Roth-Hanania, R., & Knafo, A. (2013). Concern for others in the first year of life: Theory, evidence, and avenues for research. *Child Development Perspectives*, 7(2), 126–131.
- Decarli, G., Franchin, L., Piazza, M., & Surian, L. (2020). Infants' use of motion cues in object individuation processes. *Journal of Experimental Child Psychology*, 197, 104868.
- DesChamps, T. D., Eason, A. E., & Sommerville, J. A. (2016). Infants associate praise and admonishment with fair and unfair individuals. *Infancy*, 21(4), 478–504. <https://doi.org/10.1111/infa.12117>
- Deutsch, M. (1975). Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *Journal of Social Issues*, 31(3), 137–149.
- Dunlea, J. P., & Heiphetz, L. (2020). Children's and adults' understanding of punishment and the criminal justice system. *Journal of Experimental Social Psychology*, 87, 103913.
- Dwyer, S., Huebner, B., & Hauser, M. D. (2010). The linguistic analogy: Motivations, results, and speculations. *Topics in Cognitive Science*, 2(3), 486–510. <https://doi.org/10.1111/j.1756-8765.2009.01064.x>
- Elenbaas, L. (2019). Against unfairness: Young children's judgments about merit, equity, and equality. *Journal of Experimental Child Psychology*, 186, 73–82. <https://doi.org/10.1016/j.jecp.2019.05.009>
- Engelmann, J. M., & Tomasello, M. (2019). Children's sense of fairness as equal respect. *Trends in Cognitive Sciences*, 23(6), 454–463. <https://doi.org/10.1016/j.tics.2019.03.001>
- Enright, E. A., Gweon, H., & Sommerville, J. A. (2017). 'To the victor go the spoils': Infants expect resources to align with dominance structures. *Cognition*, 164, 8–21. <https://doi.org/10.1016/j.cognition.2017.03.008>
- Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature*, 454(7208), 1079–1083. <https://doi.org/10.1038/nature07155>
- Forgács, B., Gervain, J., Parise, E., Csibra, G., Gergely, G., Baross, J., & Király, I. (2020). Electrophysiological investigation of infants' understanding of understanding. *Developmental Cognitive Neuroscience*, 43, 100783. <https://doi.org/10.1016/j.dcn.2020.100783>
- Franchin, L., Savazzi, F., Neira-Gutierrez, I. C., & Surian, L. (2019). Toddlers map the word 'good' to helping agents, but not to fair distributors. *Journal of Child Language*, 46(1), 98–110. <https://doi.org/10.1017/S0305000918000351>
- Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources. *Developmental Science*, 14, 1012–1020. <https://doi.org/10.1111/j.1467-7687.2011.01048.x>
- Geraci, A., & Surian, L. (2023a). Intention-based evaluations of distributive actions by 4-month-olds. *Infant Behavior and Development*, 70, 101797. <https://doi.org/10.1016/j.infbeh.2022.101797>
- Geraci, A., & Surian, L. (2023b). Preverbal infants' reactions to third-party punishments and rewards delivered toward fair and unfair agents. *Journal of Experimental Child Psychology*, 226, 105574. <https://doi.org/10.1016/j.jecp.2022.105574>
- Geraci, A., Simion, F., & Surian, L. (2022). Infants' intention-based evaluations of distributive actions. *Journal of Experimental Child Psychology*, 220, 105429. <https://doi.org/10.1016/j.jecp.2022.105429>
- Geraci, A., Benavides-Varela, S., Nascimben, C., Simion, F., & Di Giorgio, E. (2024). Evaluations of aggressive chasing interactions by 7-month-old infants. *Aggressive Behavior*, 50(5), e22174. <https://doi.org/10.1002/ab.22174>
- Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences*, 7(7), 287–292. [https://doi.org/10.1016/S1364-6613\(03\)00128-1](https://doi.org/10.1016/S1364-6613(03)00128-1)
- Gill, I. K., & Sommerville, J. A. (2023). Generalizing across moral sub-domains: Infants bidirectionally link fairness and unfairness to helping and hindering. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1213409>
- Graham, J., Haidt, J., Motyl, M., Meindl, P., Iskiwitsch, C., & Mooijman, M. (2018). Moral foundations theory: On the advantages of moral pluralism over moral monism. *Atlas of Moral Psychology*, 211, 222.
- Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Spangenberg, D. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. In P. Devine, A. Plant (Eds.),

- Advances in Experimental Social Psychology* (Vol. 47, pp. 55–130). Academic Press. <https://doi.org/10.1016/b978-0-12-407236-7.00002-4>
- Gray, K., Schein, C., & Ward, A. F. (2014). The myth of harmless wrongs in moral cognition: Automatic dyadic completion from sin to suffering. *Journal of Experimental Psychology: General*, 143, 600–1616.
- Grocke, P., Rossano, P., & Tomasello, M. (2015). Procedural justice in children: Preschoolers accept unequal resource distributions if the procedure provides equal opportunities. *Journal of Experimental Child Psychology*, 140, 197–210.
- Haidt, J. & Joseph, C. (2007). The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules. In P. Carruthers, S. Laurence and S. Stich (Eds.), *The innate mind* (Vol. 3), Oxford University Press.
- Hamlin, J. K. (2013a). Moral judgment and action in preverbal infants and toddlers: Evidence for an innate moral core. *Current Directions in Psychological Science*, 22(3), 186–193. <https://doi.org/10.1177/0963721412470687>
- Hamlin, J. K. (2013b). Failed attempts to help and harm: Intention versus outcome in preverbal infants' social evaluations. *Cognition*, 128(3), 451–474.
- Hamlin, J. K. (2023). Core morality? Or merely core agents and social beings? A response to Spelke's what babies know. *Mind & Language*, 38(5), 1323–1335.
- Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, 450(7169), 557–559. <https://doi.org/10.1038/nature06288>
- Hamlin, J. K., Wynn, K., & Bloom, P. (2010). Three-month-olds show a negativity bias in their social evaluations. *Developmental Science*, 13(6), 923–929. <https://doi.org/10.1111/j.1467-7687.2010.00951.x>
- Hamlin, J. K., Ullman, T., Tenenbaum, J., Goodman, N., & Baker, C. (2013). The mentalistic basis of core social cognition: Experiments in preverbal infants and a computational model. *Developmental Science*, 16(2), 209–226. <https://doi.org/10.1111/desc.12017>
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & Tracer, D. (2005). “Economic man” in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, 28(6), 795–815.
- House, B. R., Silk, J. B., Henrich, J., Barrett, H. C., Scelza, B. A., Boyette, A. H., ... & Laurence, S. (2013). Ontogeny of prosocial behavior across diverse societies. *Proceedings of the National Academy of Sciences*, 110(36), 14586–14591.
- House, B. R., Kanngiesser, P., Barrett, H. C., Yilmaz, S., Smith, A. M., Sebastian-Enesco, C., & Silk, J. B. (2020). Social norms and cultural diversity in the development of third-party punishment. *Proceedings of the Royal Society B*, 287(1925), 20192794. <https://doi.org/10.1098/rspb.2019.2794>
- Hudac, C. M., & Sommerville, J. A. (2020). Understanding others' minds and morals: Progress and innovation of infant electrophysiology. In J. Decety (Ed.), *The social brain: A developmental perspective* (pp. 41–65). MIT Press.
- Hume, D. (1975) *An Enquiry Concerning the Principles of Morals* in *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, edited by L. A. Selby-Bigge, revised by P.H. Nidditch. Oxford: Clarendon Press.
- Huppert, E., et al. (2019). The development of children's preferences for equality and equity across 13 individualistic and collectivist cultures. *Developmental Science*, 22, e12729. <https://doi.org/10.1111/desc.12729>
- Jacobs, C., Flowers, M., Aboody, R., Maier, M., & Jara-Ettinger, J. (2022). Not just what you did, but how: Children see distributors that count as more fair than distributors who don't. *Cognition*, 225, 105128. <https://doi.org/10.1016/j.cognition.2022.105128>
- Jin, K. S., Ting, F., He, Z., & Baillargeon, R. (2024). Infants expect some degree of positive and negative reciprocity between strangers. *Nature Communications*, 15(1), 7742. <https://doi.org/10.1038/s41467-024-51982-7>
- Kampis, D., Parise, E., Csibra, G., & Kovács, Á. M. (2015). Neural signatures for sustaining object representations attributed to others in preverbal human infants. *Proceedings of the Royal Society B: Biological Sciences*, 282(1819), 20151683. <https://doi.org/10.1098/rspb.2015.1683>
- Kampis, D., Buttelmann, F., & Kovács, Á. M. (2020). Developing a theory of mind: Are infants sensitive to how other people represent the world? In J. Decety (Ed.), *The social brain: A developmental perspective* (pp. 143–160). MIT Press.
- Kanngiesser, P., Rossano, F., Zeidler, H., Haun, D., & Tomasello, M. (2019). Children's respect for ownership across diverse societies. *Developmental Psychology*, 55(11), 2286.

- Killen, M. (2018). The origins of morality: Social equality, fairness, and justice. *Philosophical Psychology*, 31(5), 767–803. <https://doi.org/10.1080/09515089.2018.1486612>
- Killen, M., Elenbaas, L., & Ruck, M. D. (2022). Developmental perspectives on social inequalities and human rights. *Human Development*, 66(4–5), 329–342.
- Kohlberg, L. (1971). From is to ought: How to commit the naturalistic fallacy and get away with it in the study of moral development. In T. Mischel (Ed.), *Psychology and genetic epistemology* (pp. 151–235). New York, NY: Academic Press. <https://doi.org/10.1016/B978-0-12-498640-4.50011-1>
- Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, 330(6012), 1830–1834.
- Krebs, D. L. (2008). Morality: An evolutionary account. *Perspectives on Psychological Science*, 3(3), 149–172.
- Laurence, S., & Margolis, E. (2024). *The building blocks of thought. A rationalist account of the origins of concepts*. Oxford University Press.
- Leslie, A. M. (1994). ToMM, ToBy, and Agency: Core architecture and domain specificity. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119–148). Cambridge University Press.
- Levine, S., Mikhail, J., & Leslie, A. M. (2018). Presumed innocent? How tacit assumptions of intentional structure shape moral judgment. *Journal of Experimental Psychology: General*, 147(11), 1728–1747. <https://doi.org/10.1037/xge0000459>
- Lucca, K., Pospisil, J., & Sommerville, J. A. (2018). Fairness informs social decision making in infancy. *PLoS One*, 13(2), e0192848. <https://doi.org/10.1371/journal.pone.0192848>
- Lucca, K., Capelier-Mourguy, A., Byers-Heinlein, K., Cirelli, L., Dal Ben, R., Frank, M. C., ... Hamlin, K. (2021). Infants' social evaluation of helpers and hinderers: A large-scale, multi-Lab, coordinated replication study. <https://doi.org/10.31234/osf.io/qhxxm>
- Luo, Y. (2011). Three-month-old infants attribute goals to a non-human agent. *Developmental Science*, 14, 453–460. <https://doi.org/10.1111/j.1467-7687.2010.00995.x>
- Mammen, M., & Paulus, M. (2023). The communicative nature of moral development: A theoretical framework on the emergence of moral reasoning in social interactions. *Cognitive Development*, 66, 101336. <https://doi.org/10.1016/j.cogdev.2023.101336>
- Margoni, F., & Surian, L. (2017). Children's intention-based moral judgments of helping agents. *Cognitive Development*, 41, 46–64. <https://doi.org/10.1016/j.cogdev.2016.12.001>
- Margoni, F., & Surian, L. (2021). Question framing effects and the processing of the moral-conventional distinction. *Philosophical Psychology*, 34(1), 76–101.
- Margoni, F., Baillargeon, R., & Surian, L. (2018). Infants distinguish between leaders and bullies. *Proceedings of the National Academy of Sciences, USA*, 115, 8835–8843. <https://doi.org/10.1073/pnas.1801677115>
- Margoni, F., Surian, L., & Baillargeon, R. (2024). The Violation-of-Expectation paradigm: A conceptual overview. *Psychological Review*, 131, 716–748. <https://doi.org/10.1037/rev0000450>
- Mascaro, O., & Csibra, G. (2012). Representation of stable social dominance relations by human infants. *Proceedings of the National Academy of Sciences*, 109(18), 6862–6867. <https://doi.org/10.1073/pnas.1113194109>
- McManus, R. M., Kleiman-Weiner, M., & Young, L. (2020). What we owe to family: The impact of special obligations on moral judgment. *Psychological Science*, 31(3), 227–242. <https://doi.org/10.1177/0956797619900321>
- Meristo, M., & Surian, L. (2013). Do infants detect indirect reciprocity? *Cognition*, 129(1), 102–113. <https://doi.org/10.1016/j.cognition.2013.06.006>
- Meristo, M., & Zeidler, H. (2022). Cross-cultural differences in early expectations about third party resource distribution. *Scientific Reports*, 12(1), 11627. <https://doi.org/10.1038/s41598-022-15766-7>
- Meristo, M., Strid, K., & Surian, L. (2016). Preverbal infants' ability to encode the outcome of distributive actions. *Infancy*, 21(3), 353–372. <https://doi.org/10.1111/inf.12124>
- Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in Cognitive Sciences*, 11(4), 143–152.
- Mikhail, J. (2022). Moral intuitionism and moral nativism. In M. Vargas & J. Doris (Eds.), *Oxford handbook of moral psychology* (pp. 364–387). Oxford University Press.
- Miller, D. (2021). Justice. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. URL = <https://plato.stanford.edu/archives/fall2021/entries/justice/>.

- Nava, E., Croci, E., & Turati, C. (2019). 'I see you sharing, thus I share with you': Indirect reciprocity in toddlers but not infants. *Palgrave Communications*, 5, 66. <https://doi.org/10.1057/s41599-019-0268-z>
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308(5719), 255–258. <https://doi.org/10.1126/science.1107621>
- Osler, G., Franchin, L., Guglielmetti, G., Calzolari, S., Micciolo, R., & Surian, L. (2024). The implicit perception of harm following moral violations in autism. *Thinking & Reasoning*, 30(2), 379–393.
- Paulus, M. (2020). How do young children become moral agents? A developmental perspective. In J. Decety (Ed.), *The social brain: A developmental perspective* (pp. 41–65). MIT Press.
- Pellizzoni, S., Siegal, M., & Surian, L. (2010). The contact principle and utilitarian moral judgments in young children. *Developmental Science*, 13(2), 265–270. <https://doi.org/10.1111/j.1467-7687.2009.00851.x>
- Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification*. American Psychological Association.
- Piaget, J. (1932/1965). *The moral judgment of the child*. Free Press.
- Piazza, J., & Sousa, P. (2016). When injustice is at stake, moral judgements are not parochial. *Proceedings of the Royal Society b: Biological Sciences*, 283(1823), 20152037.
- Premack, D., & Premack, A. J. (1997). Infants attribute value± to the goal-directed actions of self-propelled objects. *Journal of Cognitive Neuroscience*, 9(6), 848–856. <https://doi.org/10.1162/jocn.1997.9.6.848>
- Prinz, J. (2006). Is morality innate? In W. Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol 1). Cambridge, MA: MIT Press (pp. 367–406).
- Pun, A., Birch, S. A., & Baron, A. S. (2016). Infants use relative numerical group size to infer social dominance. *Proceedings of the National Academy of Sciences*, 113(9), 2376–2381.
- Rawls, J. (1958). Justice as fairness. *The Philosophical Review*, 67(2), 164–194.
- Rawls, J. (1971). *A theory of justice*. Harvard University Press.
- Revenu, B., & Csibra, G. (2023). The missing link between core knowledge and language: Review of Elizabeth Spelke's What babies know, volume 1 (2022). *Mind & Language*, 38(5), 1314–1322.
- Rochat, P., Dias, M. D. G., Liping, G., Broesch, T., Passos-Ferreira, C., Winning, A., & Berg, B. (2009). Fairness in distributive justice by 3- and 5-year-olds across seven cultures. *Journal of Cross-Cultural Psychology*, 40, 416–442. <https://doi.org/10.1177/0022022109332844>
- Ruffman, T., Taumoepeau, M., & Perkins, C. (2012). Statistical learning as a basis for social understanding in children. *British Journal of Developmental Psychology*, 30(1), 87–104. <https://doi.org/10.1111/j.2044-835X.2011.02045.x>
- Schmidt, M. F., & Sommerville, J. A. (2011). Fairness expectations and altruistic sharing in 15-month-old human infants. *PLoS ONE*, 6(10), e23223. <https://doi.org/10.1371/journal.pone.0023223>
- Scott, R. M., & Baillargeon, R. (2017). Early false-belief understanding. *Trends in Cognitive Sciences*, 21(4), 237–249. <https://doi.org/10.1016/j.tics.2017.01.012>
- Sen, A. (2009). *The idea of justice*. Harvard University Press.
- Shaw, A., & Olson, K. R. (2014). Fairness as an aversion to partiality: The development of procedural justice. *Journal of Experimental Child Psychology*, 119, 40–53.
- Sheskin, M., Chevallier, C., Lambert, S., & Baumard, N. (2014). Life-history theory explains childhood moral development. *Trends in Cognitive Sciences*, 18, 613–615. <https://doi.org/10.1016/j.tics.2014.08.004>
- Shweder, R. A. (2004). Moral realism without the ethnocentrism: Is it just a list of empty truisms? In András Sajó (ed.), *Human rights with modesty: The problem of universalism* (pp. 65–102). The Netherlands: Koninklijke Brill.
- Sloane, S., Baillargeon, R., & Premack, D. (2012). Do infants have a sense of fairness? *Psychological Science*, 23(2), 196–204. <https://doi.org/10.1177/0956797611422072>
- Smetana, J. G. (1983). Social cognitive development: Domain distinctions and coordination. *Developmental Review*, 3, 131–147. [https://doi.org/10.1016/0273-2297\(83\)90027-8](https://doi.org/10.1016/0273-2297(83)90027-8)
- Smetana, J. G. (1984). Toddlers' interactions regarding moral and conventional transgressions. *Child Development*, 55, 1767–1776. <https://doi.org/10.2307/1129924>
- Sodian, B., Licata, M., Kristen-Antonow, S., Paulus, M., Killen, M., & Woodward, A. (2016). Understanding of goals, beliefs, and desires predicts morally relevant theory of mind: A longitudinal investigation. *Child Development*, 87(4), 1221–1232. <https://doi.org/10.1111/cdev.12533>

- Sodian, B., Kristen-Antonow, S., & Kloo, D. (2020). How does children's theory of mind become explicit? A review of longitudinal findings. *Child Development Perspectives*, 14(3), 171–177. <https://doi.org/10.1111/cdep.12381>
- Sommerville, J. A. (2018). Infants' understanding of distributive fairness as a test case for identifying the extents and limits of infants' sociomoral cognition and behavior. *Child Development Perspectives*, 12, 141–145. <https://doi.org/10.1111/cdep.12283>
- Sommerville, J. (2023). Developing an early awareness of fairness. In M. Killen & J. S. Smetana (eds.), *Handbook of moral development* (3rd ed.). Routledge.
- Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, 96, B1–B11. <https://doi.org/10.1016/j.cognition.2004.07.004>
- Sommerville, J. A., Schmidt, M. F., Yun, J. E., & Burns, M. (2013). The development of fairness expectations and prosocial behavior in the second year of life. *Infancy*, 18(1), 40–66. <https://doi.org/10.1111/j.1532-7078.2012.00129.x>
- Southgate, V., & Vernetti, A. (2014). Belief-based action prediction in preverbal infants. *Cognition*, 130(1), 1–10. <https://doi.org/10.1016/j.cognition.2013.08.008>
- Spelke, E. S. (2022). *What babies know: Core knowledge and composition* (Vol. 1). Oxford University Press.
- Spelke, E. S. (2023). Core knowledge, language learning, and the origins of morality and pedagogy: Reply to reviews of *What babies know*. *Mind and Language*, 38, 1336–1350.
- Sperber, D., & Baumard, N. (2012). Moral reputation: An evolutionary and cognitive perspective. *Mind & Language*, 27(5), 495–518.
- Stahl, A. E., & Feigenson, L. (2019). Violations of core knowledge shape early learning. *Topics in Cognitive Science*, 11(1), 136–153.
- Stavans, M., & Baillargeon, R. (2019). Infants expect leaders to right wrongs. *Proceedings of the National Academy of Sciences*, 116(33), 16292–16301.
- Stowe, L. M., Peretz-Lange, R., & Blake, P. R. (2022). Children consider procedures, outcomes, and emotions when judging the fairness of inequality. *Frontiers in Psychology*, 13, 815901.
- Strid, K., & Meristo, M. (2020). Infants consider the distributor's intentions in resource allocation. *Frontiers in Psychology*, 11, 596213. <https://doi.org/10.3389/fpsyg.2020.596213>
- Surian, L., & Caldi, S. (2010). Infants' individuation of agents and inert objects. *Developmental Science*, 13(1), 143–150. <https://doi.org/10.1111/j.1467-7687.2009.00873.x>
- Surian, L., & Franchin, L. (2017a). Toddlers selectively help fair agents. *Frontiers in Psychology*, 8, 944. <https://doi.org/10.3389/fpsyg.2017.00944>
- Surian, L., & Franchin, L. (2017b). Infants reason about deserving agents: A test with distributive actions. *Cognitive Development*, 44, 49–56. <https://doi.org/10.1016/j.cogdev.2017.08.009>
- Surian, L., & Margoni, F. (2020). First steps toward an understanding of procedural fairness. *Developmental Science*, 23(5), e12939. <https://doi.org/10.1111/desc.12939>
- Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, 18(7), 580–586. <https://doi.org/10.1111/j.1467-9280.2007.01943.x>
- Surian, L., Ueno, M., Itakura, S., & Meristo, M. (2018). Do infants attribute moral traits? Fourteen-month-olds' expectations of fairness are affected by agents' antisocial actions. *Frontiers in Psychology*, 9, 1649. <https://doi.org/10.3389/fpsyg.2018.01649>
- Tan, E., & Hamlin, J. K. (2022). Infants' neural responses to helping and hindering scenarios. *Developmental Cognitive Neuroscience*, 54, 101095. <https://doi.org/10.1016/j.dcn.2022.101095>
- Tan, E., Mikami, A. Y., & Hamlin, J. K. (2018). Do infant sociomoral evaluation and action studies predict preschool social and behavioral adjustment? *Journal of Experimental Child Psychology*, 176, 39–54. <https://doi.org/10.1016/j.jecp.2018.07.003>
- Tatone, D., Hernik, M., & Csibra, G. (2019). Minimal cues of possession transfer compel infants to ascribe the goal of giving. *Open Mind*, 3, 31–40. https://doi.org/10.1162/opmi_a_00024
- Thomsen, L., Frankenhuys, W. E., Ingold-Smith, M., & Carey, S. (2011). Big and mighty: Preverbal infants mentally represent social dominance. *Science*, 331, 477–480. <https://doi.org/10.1126/science.1199198>
- Ting, F., & Baillargeon, R. (2021). Toddlers draw broad negative inferences from wrongdoers' moral violations. *Proceedings of the National Academy of Sciences*, 118(39), e2109045118. <https://doi.org/10.1073/pnas.2109045118>

- Ting, F., Buyukozur Dawkins, M., Stavans, M., & Baillargeon, R. (2020). Principles and concepts in early moral cognition. In J. Decety (Ed.), *The social brain: A developmental perspective* (pp. 41–65). MIT Press.
- Ting, F. (2020). *Early reasoning about moral obligations and rights*. Doctoral dissertation, University of Illinois at Urbana-Champaign.
- Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two key steps in the evolution of human cooperation: The interdependence hypothesis. *Current Anthropology*, 53(6), 673–692. <https://doi.org/10.1086/668207>
- Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge University Press.
- Vaish, A. & Tomasello, M. (2023). The early ontogeny of human cooperation and morality. In M. Killen & J. S. Smetana (eds.), *Handbook of moral development*. 3rd Edition. Routledge.
- Wang, Y., & Henderson, A. M. (2018). Just rewards: 17-month-old infants expect agents to take resources according to the principles of distributive justice. *Journal of Experimental Child Psychology*, 172, 25–40. <https://doi.org/10.1016/j.jecp.2018.02.008>
- Woo, B. M., Steckler, C. M., Le, D. T., & Hamlin, J. K. (2017). Social evaluation of intentional, truly accidental, and negligently accidental helpers and harmers by 10-month-old infants. *Cognition*, 168, 154–163. <https://doi.org/10.1016/j.cognition.2017.06.029>
- Woo, B. M., & Hamlin, J. K. (2023). Evidence for an early-emerging moral core. In M. Killen & J. S. Smetana (eds.), *Handbook of moral development*. 3rd Edition. Routledge.
- Woo, B. M., & Spelke, E. S. (2022). Toddlers' social evaluations of agents who act on false beliefs. *Developmental Science*, e13314. <https://doi.org/10.1111/desc.13314>
- Woo, B. M., Tan, E., & Hamlin, J. K. (2022). Human morality is based on an early-emerging moral core. *Annual Review of Developmental Psychology*, 4(1), 41–61. <https://doi.org/10.1146/annurev-devpsy-121020-023312>
- Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, 69(1), 1–34. [https://doi.org/10.1016/s0010-0277\(98\)00058-4](https://doi.org/10.1016/s0010-0277(98)00058-4)
- Wynn, K., & Bloom, P. (2014). The moral baby. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 435–453). Psychology Press.
- Yucel, M., Drell, M. B., Jaswal, V. K., & Vaish, A. (2022). Young children do not perceive distributional fairness as a moral norm. *Developmental Psychology*, 58(6), 1103–1113. <https://doi.org/10.1037/dev0001349>
- Ziv, T., & Sommerville, J. A. (2017). Developmental differences in infants' fairness expectations from 6 to 15 months of age. *Child Development*, 88(6), 1930–1951. <https://doi.org/10.1111/cdev.12674>
- Ziv, T., Whiteman, J. D., & Sommerville, J. A. (2021). Toddlers' interventions toward fair and unfair individuals. *Cognition*, 214, 104781. <https://doi.org/10.1016/j.cognition.2021.104781>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Luca Surian is Full Professor of Developmental Psychology at the University of Trento. He has worked as a Research Scientist for the Medical Research Council, Cognitive Development Unit (London) and has published extensively on several areas of psychological development, with a special interest on infants' theory of mind, object individuation processes, moral judgment, verbal communication and pragmatics. His research uses implicit and explicit behavioral measures to cast light on typical and atypical development, such as autism.

Eugenio Parise Ph.D., is Associate Professor at the University of Trento, Center for Mind/Brain Sciences, Italy. Formerly a Lecturer at Lancaster University, Department of Psychology, UK, his research interest focuses on infants' sensitivity to communicative signals, how infants perceive and process signals such as direct eye contact and infant-directed speech, and how communication drives infant cognition. He uses different research methods such as looking time studies, eye tracking, EEG - including ERPs, Time-Frequency analysis and Frequency Tagging, fNIRS.

Alessandra Geraci Ph.D., is Assistant Professor at the University of Catania (Italy). Her current projects concern the origin and development of psychological reasoning, socio-moral cognition, and social causality. She is exploring the origins of the Social Mind in newborns, as well as in young infants and children. She uses different research methods, looking time studies, and eye tracking techniques.