The developmental foundations of human fairness

Katherine McAuliffe[1]*, Peter R. Blake[2], Nikolaus Steinbeis[3,4] and Felix Warneken[5]

New behavioural and neuroscientific evidence on the development of fairness behaviours demonstrates that the signatures of human fairness can be traced into childhood. Children make sacrifices for fairness (1) when they have less than others, (2) when others have been unfair and (3) when they have more than others. The latter two responses mark a critical departure from what is observed in other species because they enable fairness to be upheld even when doing so goes against self-interest. This new work can be fruitfully combined with insights from cognitive neuroscience to understand the mechanisms of developmental change.

Fairness is a hallmark of our species's ability to maintain cooperative relationships with large numbers of unrelated — and often unfamiliar — others. However, there is much debate over the psychological foundations of this sense of fairness. Which aspects of fairness are engrained in our biology and which depend on learned social norms? What are the psychological mechanisms and motivations that give rise to human fairness? Why does fairness appear to drive decisions in some contexts but not others? To answer these questions, it is critical to understand how fairness emerges in child development, as studies of adults alone do not allow us to differentiate between psychological processes that are acquired through socialization and those that have deeper biological roots. Indeed, a complete understanding of any behaviour requires a description of its developmental course and an examination of what causes behavioural change over the lifespan[1–3].

Studies of child development can provide unique insight into the psychology of fairness in important ways: first, developmental data can help to identify the aspects of fairness that are foundational and those that are more malleable. Second, charting how fairness behaviour changes across development can shed light on the specific cognitive mechanisms that enable its emergence and expression. Finally, combining insights from development with work examining how fairness is instantiated at the neural level can help us to understand the psychological foundations on which the human sense of fairness is built.

In this Review, we appraise and integrate recent developmental and neuroscientific evidence on fairness. The resulting neurodevelopmental perspective provides novel insights into the foundations of fairness in our species. First, we argue that the signatures of our uniquely human sense of fairness are present in childhood. We then summarize the current state of knowledge for neural mechanisms that support fairness and highlight recent work on the developmental neuroscience of fairness. Last, we identify key neural processes that change during childhood and propose how the developmental integration of these systems can explain the temporal emergence of the human sense of fairness.

The signatures of human fairness

Fairness is a complex concept that is applied across many social contexts. A central focus for both theoretical and empirical work

on fairness has been the principles governing the allocation of resources, so-called distributive justice. Specifically, multiple experimental paradigms have been developed to explore how both adults and children divide up valuable resources and respond to unequal resource allocations (for example, see Box 1). Equality acts as an objective benchmark in investigations of fairness[4,5] and provides a foundation for mature conceptions of fairness[6]. Understanding how people react to unequal distributions, all else equal (that is, in the absence of richer contextual information like neediness and deservedness[7]), provides the key starting point for unlocking the psychology of fairness in humans. Indeed, this focus on equality as an objective benchmark of fairness has led to an explosion of studies across disciplines, including psychology, economics and anthropology[8–10]. In our review of children's developing fairness behaviour, we adopt this same focus on equality and, more specifically, on children's responses to unequal outcomes. Additionally, we identify the motivations and mechanisms that lead to equal outcomes across development.

One striking feature of adult fairness is that people are sometimes willing to pay a cost to avoid inequality. This has been shown in four main contexts (Fig. 1). First, people pay to achieve equality when they receive less than others (disadvantageous inequity aversion[8,9,11]). Second, adults not only make sacrifices to be fair themselves, they also punish those who are unfair. This punishment comes in two forms, as second-party punishment, where the punisher has been directly affected by unfair behaviour[12,13] and as third-party punishment, where the punisher witnesses unfairness happening to others but is not directly affected by it[14,15]. Importantly, in both forms of punishment, the punisher punishes unfairness in others at personal cost or, at least, to no immediate personal benefit, thus often overriding self-interest to enforce fairness. Finally, people reject unfair distributions even when they receive more than others (advantageous inequity aversion[8,9,11]), demonstrating costly conformity to a fairness norm.

We organize our review of children's emerging fairness behaviour around these four signatures of human fairness. This structure is grounded in behavioural game theory[10] and provides a taxonomy for exploring children's decisions about fairness in four key contexts in which they must reconcile their own self-interest with a desire for fairness. This taxonomy also represents, we argue, a developmental

[1]Department of Psychology, Boston College, Chestnut Hill, Massachusetts 02467, USA. [2]Department of Psychological and Brain Sciences, Boston University, Boston, Massachusetts 02215, USA. [3]Department of Developmental Psychology, Leiden University, Leiden 2333AK, The Netherlands. [4]Department of Social Neuroscience, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig 04103, Germany. [5]Department of Psychology, Harvard University, Cambridge, Massachusetts 02138, USA. *e-mail: katherine.mcauliffe.2@bc.edu

Dictator Game
Player One is given an endowment of a resource and can give some or all of their endowment to Player Two. Player Two is a passive recipient.

Inequity Game
An allocation is distributed between Players One and Two by a third party (for example, an experimenter). Player One responds to the allocation while Player Two is a passive recipient. If Player One accepts the allocation, both players receive their allocated payoffs. If Player One rejects the allocation, neither player receives a payoff.

Third-Party Punishment Game
A variation of the Dictator Game (see above) in which a third player (Player Three) observes Player One's donation to Player Two. Player Three can then intervene to reduce Player One's earnings, which constitutes third-party punishment.

Ultimatum Game
Player One is given an endowment of a resource and can offer any portion of their endowment to Player Two. Player Two can either accept or reject Player One's offer. If Player Two accepts, they both receive their allocated portions, as proposed by Player One. If Player Two rejects, neither player receives a payoff. Rejections in this context constitute second-party punishment.

progression that is relevant for theories of fairness. Most broadly, children show an aversion to receiving less than others first — a focus on unfairness to self — and only later engage in decisions that correct or prevent unfairness to others. Theoretically, this progression marks a key shift towards understanding fairness as a norm that should be enforced, an important marker of a uniquely human sense of fairness.

Of the four contexts described above, only behaviour broadly consistent with disadvantageous inequity aversion has been found in non-human species[16–18]. These data suggest that an aversion to getting less than others — or, at least, to getting less than you could be getting — has deep biological roots. However, there is no evidence that non-human animals enforce fairness through second- and third-party punishment[19–21] and virtually no evidence that they reject advantageous allocations to prevent others from getting less than them[18]. Punishment and advantageous inequity aversion thus constitute clear signatures of the uniquely human sense of fairness. These reactions mark a critical departure from what is observed outside of humans because they enable fairness norms to be upheld even when doing so is not motivated by self-interest. In particular, the enforcement of norms by unaffected third parties is theorized to be a key means of stabilizing cooperative behaviour in humans[22]. Sanctioning norm violators in this context depends on the belief that 'others ought to be fair generally' not just that 'others ought to be fair to me'. In thinking about the origins of this behaviour, a critical question emerges: namely, when do children shift from being the ones sanctioned for fairness norm violations to the ones sanctioning others? The answer to this question bears on deep theoretical issues about the origins of cooperation in development and, in particular, the origins of the uniquely human desire to promote good behaviour in others.

In recent years, research has begun to examine when, during development, children begin to respond negatively to unequal resource distributions and, critically, when they will pay to avoid them (Box 2). In what follows, we summarize evidence showing that the four signatures of human fairness are present in childhood, with disadvantageous inequity aversion emerging early, a willingness to punish unfairness in others (second- and third-party punishment) emerging next and advantageous inequity aversion appearing last. We then ground this description of behavioural phenomena in a framework of potential neurocognitive mechanisms underlying the emergence of these fairness behaviours.

Children show the signatures of human fairness
Research on the development of fairness has undergone a shift in recent years from a focus on children's explicit verbal reasoning about hypothetical situations to a focus on their behavioural responses in real situations (see Box 2). Drawing on paradigms from behavioural economics (Box 1), new evidence shows that children engage in sophisticated fairness behaviours much younger than previously believed.

Children avoid getting less than others (disadvantageous inequity aversion). Several recent studies have found that infants are surprised when witnessing resources distributed unequally for no apparent reason[23–25], and by three years of age children will state explicitly that equality is a norm that should be followed in distribution tasks[26]. However, when young children face inequality themselves, they are measurably upset at receiving less, but quite happy to receive more[27]. By four years of age, children will pay a cost to prevent a peer from getting more, showing the first signs of disadvantageous inequity aversion. In one task, the Inequity Game (Box 1), children could accept or reject different allocations of sweets[28] (Box 2). Children as young as four years of age rejected allocations that placed them at a disadvantage relative to a peer, sacrificing a small amount of candy to prevent their partner from receiving more. Nevertheless, these rejections were not automatic — children took longer to make their decisions when facing a disadvantage compared with when they faced equal trials. This finding suggests that children experience a conflict between self-interest and fairness to the self when confronted with disadvantageous inequity. Results from the Inequity Game and other experimental paradigms show that children's aversion to disadvantageous outcomes tends to grow stronger with age[28,29]. Interestingly, recent evidence suggests that children are more tolerant of a disadvantage they create themselves, a prosocial act[30,31].

Subsequent work has shown that disadvantageous inequity aversion is even observed in non-social contexts (that is, when a child receives the lesser of two alternative payoffs), although the response is stronger in social situations when the larger payoff goes to a peer[32]. Additional work has shown that children's rejections of disadvantageous allocations are motivated by spite — the desire to deprive a peer of the better payoff[33]. Most recently, this task was conducted across seven societies to characterize cross-cultural variation in the development of disadvantageous inequity aversion. Results showed that children across all societies rejected disadvantageous allocations, providing evidence that disadvantageous inequity aversion may be a universal response in humans[34].

Children punish those who have been unfair to them (second-party punishment). While in the Inequity Game, children make decisions about pre-existing distributions, in bargaining tasks such as the Ultimatum Game (Box 1) these distributions are actual offers made by a peer. In this context, disadvantageous (selfish) offers can be viewed as an intentional slight and the responder has the opportunity to punish the proposer by rejecting all of the rewards. Research consistently shows that with age, children become more likely to reject unfair offers. While few studies have tested young children, existing studies show that children at the ages of 4–5 years tend to accept everything, even highly unfair offers[35–37]. By contrast, starting around 6–8 years of age, children begin to systematically
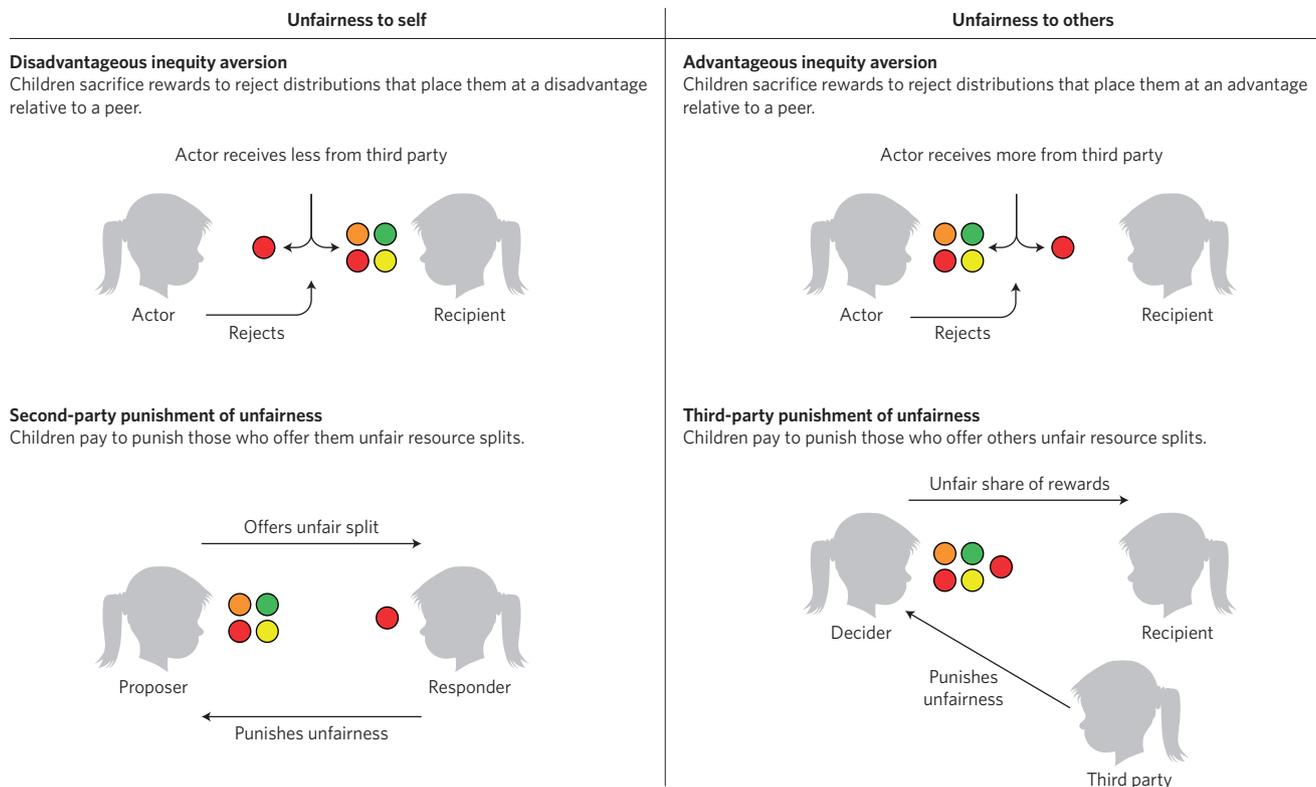
Unfairness to self	Unfairness to others

Disadvantageous inequity aversion
Children sacrifice rewards to reject distributions that place them at a disadvantage relative to a peer.

Advantageous inequity aversion
Children sacrifice rewards to reject distributions that place them at an advantage relative to a peer.

Second-party punishment of unfairness
Children pay to punish those who offer them unfair resource splits.

Third-party punishment of unfairness
Children pay to punish those who offer others unfair resource splits.

Figure 1 | Diagram of four classes of fairness behaviour in children.

reject unfair offers in which they would obtain less than half of the resource[37–43], a response that becomes even stronger with age[38,39,42,44,45] (but see ref. [43]). These studies point to important age differences, with the first clear signs of children being willing to pay a cost to punish unfair treatment at around 6–8 years of age.

When in the role of the proposer in the Ultimatum Game, older, but not younger children, make mostly fair offers, perhaps because they realize that the responder will not accept a disadvantageous proposal. To assess whether this change is truly strategic or represents a shift to more generous behaviour, researchers have compared children's offers in Ultimatum Games with their offers in Dictator Games (Box 1). If children are able to anticipate that the recipient will punish them for unfair offers, they should offer the recipient significantly more in Ultimatum Games (where the recipient can punish) than Dictator Games (where the recipient has no such power). Studies show that in the age range of 6 to 8 years, children offer significantly more as proposers in Ultimatum Games than in Dictator Games[37,40,43–45]. Few studies have tested younger children, but existing studies find that younger children do not offer significantly more in the Ultimatum Game than the Dictator Game[36,37], suggesting that more strategic behaviour emerges with age.

Taken together, both children's behaviour as responders and proposers show that starting around 6 to 8 years of age, children engage in and expect second-party punishment: they will pay a cost to prevent unfair treatment towards themselves and anticipate that others will engage in this behaviour as well.

Children punish those who have been unfair to others (third-party punishment). Costly third-party punishment of unfairness requires that people combine a reaction to unfairness with a willingness to intervene against bad behaviour, even when they are unaffected bystanders. Developmental data from infancy suggest that the building blocks for third-party punishment can be found early in ontogeny: infants and toddlers dislike antisocial agents

who, for example, slam a box shut that another agent is trying to open and like individuals who punish these antisocial agents[46,47]. By mid-childhood, children will punish individuals for violating fairness norms, marking the emergence of costly third-party punishment[38,48–50]. For instance, when children are presented with scenarios in which they learned of a selfish child who had refused to share candies with a partner, six-year-olds will sacrifice their own resources to take candies away from the selfish individual[48]. Another study testing older children and adolescents showed that 8-, 12- and 15-year-olds punished others for unfair resource divisions, with the degree of punishment mapping onto the degree of unfairness[38].

Children's incipient willingness to engage in third-party punishment is part of an important shift away from a myopic focus on promoting fairness for self to a more generalized focus on promoting fairness for others. This shift requires an integration of a willingness to punish at a cost to oneself with recognition of fairness as a norm that must be followed by all. Importantly, this behaviour places children in a more active role as enforcers of cooperative norms. The emergence of costly third-party punishment in children suggests that the motivation and skills needed to promote cooperation in others are present in mid-childhood.

Children sometimes avoid getting more than others (advantageous inequity aversion). Although young children refuse inequity when they receive less than others, they seem perfectly willing — even eager— to receive more than others[28,51]. However, by 8 years of age children begin to reject advantageous allocations in the Inequity Game, incurring a relatively high cost to ensure that a peer does not receive less than them[28,32,34]. Children in this age group take longer to decide what to do when faced with an advantageous allocation compared with an equal allocation, indicating that children experience a conflict when faced with this decision[28]. Although the nature of the conflict is unclear, the large reward to oneself creates a tension between self-interest and a desire either to enact fairness norms or,

Box 2 | Testing fairness behaviour in children.

Research on children's understanding of fairness has a long history in developmental psychology. Early methods focused on interviews and hypothetical scenarios, which have the advantage of removing immediate self-interest and allowing children to express their judgments about fair allocations. However, these methods do not assess children's fairness behaviour when actual resources are at stake. A recent integration of methods from behavioural economics into developmental research has allowed for tests of children's fairness behaviour when achieving fair outcomes is costly. Often the decisions children face are simplified to reduce the need for considering multiple alternative distributions and to contrast specific choices in order to reveal children's motivations and priorities. There are two main classes of paradigm that have been used to study fairness behaviour in children: paradigms in which children can select equality over other options and tasks that measure children's responses to unequal allocations.

Choosing equality over other options
One common choice paradigm for children involves selecting one of two possible distributions of resources between the child and a peer[29,91]. For example, children can choose either one candy each (1–1) or two for the self and zero for the peer (2–0). Children receive other choice sets as well, typically a prosocial choice (1–1 versus 1–0) and an envy choice (1–1 versus 1–2). Across these different comparisons, children choose the equal option increasingly more often with age, which has been interpreted as evidence for the emergence of an egalitarian motive. Variations on these paradigms have also allowed researchers to identify a period in development during which children actively seek an advantage over peers, even when doing so is costly, as opposed to simply avoiding a disadvantage[51]. Children were presented with two distribution options, one of which was equal (2–2) and one of which provided them with an advantage (1–0). Researchers found that 5–6 year olds were willing to accept a lower payoff overall to achieve a relative advantage over their peer.

Measuring responses to unfairness
The choice of an equal outcome in the forced choice tasks could be supported by several mechanisms: altruistic motives (everyone should get at least one), egalitarian motives (everyone should get the same) or advantageous inequity aversion (I should not get more). A different paradigm is needed to examine the specific mechanisms that underlie the choice of equality. The Inequity Game achieves this by making the equal outcome both costly and not altruistic. Here, children face a forced choice decision that is either equal (0–0) or unequal. Each pair of children tested receives only one form of inequality: disadvantageous inequity (DI), 1–4; or advantageous inequity (AI), 4–1. Children can choose the equal option by rejecting the unequal allocation, so that both children receive nothing, or accept the inequality. Presenting the choices in this way isolates the mechanism at work. Rejecting either DI or AI denies rewards to the peer and is thus not altruistic. Rejections also deny rewards to the self and minimize the total rewards. By comparing children's rejections of DI or AI with rejections of equal allocations (1–1), researchers have identified unique mechanisms that motivate children to choose equal outcomes: disadvantageous inequity aversion ('I should not get less than you') and advantageous inequity aversion ('I should not get more than you').

While the Inequity Game isolates children's responses to unequal resource distributions, it does not test how children react to intentional unfairness. To test these reactions, the Ultimatum Game and Third-Party Punishment Game (Box 1) have been adapted for children. For instance, in one commonly used version of the Ultimatum Game, children are told that they will be paired with a partner who will either accept or reject their offer of resources. They are then presented with all possible resource divisions and asked to select what they would like to do. Using the same method, children can act as responders and either accept or reject offers made by a partner. This method has been successfully implemented in the scanner with children and adolescents, allowing researchers to measure brain activity associated with different fairness behaviours (see figure, panel c).
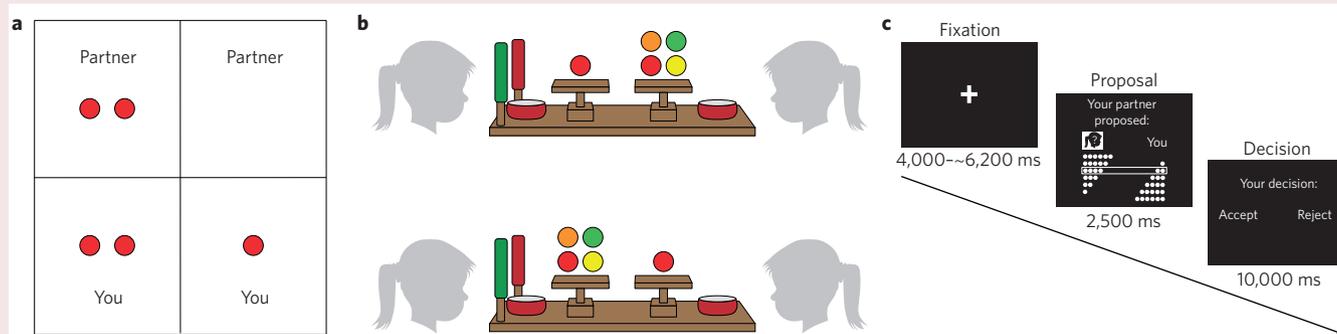
Three examples of paradigms used to test fairness behaviour in children. a, A choice task in which the child has a choice between the equal option, two for self and two for partner, or the advantageous — but lower overall — option, one for self and zero for partner. **b,** The Inequity Game used to test children's responses to disadvantageous (upper) and advantageous (lower) unfairness. In this task the actor has access to two handles and decides whether to accept (green handle) or reject (red handle) different allocations. **c,** The Ultimatum Game as presented to children in the scanner.

at least in face-to-face contexts like the Inequity Game, a desire to appear fair to the partner. Additionally, unlike rejections of disadvantageous inequality, which occur both in social and non-social contexts (that is, when the child is not paired with a partner), children exclusively reject advantageous inequity when they are paired with a social partner, suggesting that advantageous inequity aversion is an inherently social response[32].

Other experimental paradigms have found a similar shift at this age in which children now pay a cost to reduce their advantage over others[29,52]. Interestingly, one recent cross-cultural study suggests that this behaviour may not be universal during childhood[34]. Experiments with the Inequity Game found advantageous inequity aversion in late childhood in the United States, Canada and Uganda, but not in India, Mexico, Peru and Senegal. This study offers further
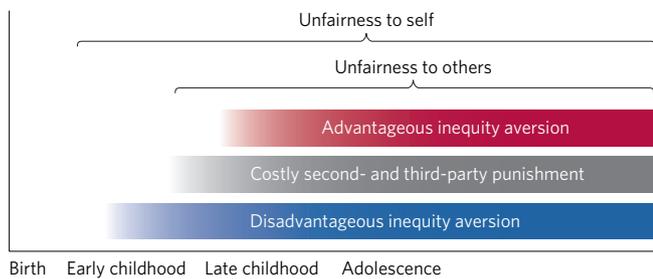
Figure 2 | The emergence of fairness behaviours in childhood.

evidence that advantageous inequity aversion is tied to local norms of fairness and expectations of child behaviour that vary by culture.

These findings show that an important developmental shift occurs in late childhood: namely, a focus on ensuring that 'I get treated well' is matched by a desire to ensure that 'others also get treated well', resulting in a more generalized preference for fairness. Existing cross-cultural comparisons suggest that the exact time at which this developmental shift occurs may vary, perhaps due to the acquisition of local norms.

Synthesis of the emergence of fairness behaviours in children.
Children show the signatures of our uniquely human sense of fairness: not only do they react to unfairness when they get less than others, they also reject unfairness when it is advantageous and pay to punish others for unfair behaviour. These data paint an interesting picture with regards to the time course of changes in fairness-related behaviours, with concerns for fairness to self emerging first, a willingness to punish unfairness in others emerging next and, finally, a willingness to reject advantageous unfairness emerging last (Fig. 2). This trajectory provides a clear example of why developmental studies are so useful: they demonstrate that these behaviours are dissociated across development, despite appearing to be united in adulthood. In doing so, this evidence provides a theoretical motivation for considering potentially distinct neural mechanisms that underlie these behaviours. In the next section, we review evidence for the cognitive and neural mechanisms that underlie fairness behaviours in children. We suggest candidate neurocognitive processes that may help shift children from an early-emerging self-regarding form of fairness to a later-emerging willingness to both conform to and enforce fairness norms even at cost to the self.

Mechanisms supporting the emergence of fairness
Although few developmental studies exist on neurocognitive mechanisms of fairness specifically, much is known about brain development more generally (Box 3). Further, research on neurocognitive mechanisms of fairness in adults can inform our understanding of what neural changes are required to reach a mature sense of fairness. In this section, we highlight the role of neurocognitive mechanisms — both common and unique — for each of our four signature fairness behaviours.

Shared components of fairness. The four signatures of human fairness vary along different psychological dimensions but also have certain key elements in common. Children must first engage in social comparison to note the difference in quantities allocated between players and then judge the difference to be unfair or undesirable. While this ability appears to already be in place in infancy[23–25], it does not in and of itself explain the emergence of fair behaviours. A critical aspect of fair behaviour is the willingness to incur a cost to create equality. Such costs create a conflict between self-interest and a fair outcome. Resolving this conflict and implementing the fair outcome thus also requires behavioural control.

Box 3 | Development of functional networks in childhood.

Studies of brain development have undergone a shift from charting the maturational processes of single regions[92–94] to characterizing structural and functional neural networks[95–97]. The prominent use of resting-state functional connectivity magnetic resonance imaging (rs-fcMRI) could show that connectivity patterns in well-characterized networks such as the default-mode network (that is, medial prefrontal cortex, precuneus, and temporo-parietal junction), the salience network (that is, ACC and anterior insula) and a central executive network (that is, dorsolateral prefrontal cortex and lateral parietal cortices) shift and change drastically throughout childhood[96]. These changes include a decrease in short-range and increase in long-range connectivity[98], a shift in functional organization from local to global networks[99,100] and increased integration between prominent functional networks from childhood through adulthood[101]. The functional consequences of these changes have been linked to robust developmental changes in higher cognitive functions such as behavioural control[102,103].

Network properties constitute a fundamental organizing principle of cognitive brain function, which is subject to considerable changes throughout development and in turn accounts for observed changes in cognitive functions. We apply this logic to understanding developmental changes in fairness behaviour. Reaching fair decisions is a vastly complex cognitive feat, requiring multiple operations. We argue that this process is subserved by the integration of signals from distal brain regions into a neural signal that guides the decision-making process. This has been shown repeatedly in the context of fair decisions in adults[56,63,65,69] (see also ref. [76] for a similar approach to mechanisms of norm enforcement). We argue that the development of neural pathways connecting distal brain regions with each other constitutes the key mechanism supporting developmental transitions in fair decision-making.

It is well documented that behavioural control improves considerably throughout childhood, becoming increasingly proactive, flexible and self-guided[53,54]. Behavioural control therefore constitutes a primary candidate to account for developmental shifts in fairness behaviours as the costs to create fair outcomes increase.

Research on neurocognitive mechanisms of behavioural responses to unfairness in adults has identified a network of brain regions that is activated during fairness decisions. These regions can be tied to the cognitive processes shared across the fairness behaviours and includes the anterior insula, the anterior cingulate cortex (ACC), ventromedial prefrontal cortex (VMPFC) as well as dorso- and ventrolateral prefrontal cortices (DLPFC and VLPFC, respectively) and striatal brain regions[55–57] (see Fig. 3). In the context of fair decisions, these brain regions subserve specific functions, such as social comparison (ventral striatum), sensitivity to fairness and emotional responses to inequality (anterior insula), conflict between following norms and one's own desires (ACC), complying with social norms (DLPFC) and implementing fair decisions (VMPFC). Thus, fairness behaviours are undergirded by a whole network of brain regions, supporting a multitude of cognitive and affective processes that interact with one another to bring about fair decisions[56–58]. This outline of the neural systems involved provides a starting point for identifying the role of specific mechanisms needed to implement the four signature fairness behaviours and how they change with age.

The role of negative emotions in creating fairness for oneself.
Developmental studies report that preschool children display negative emotional reactions when they receive less than a peer[27].
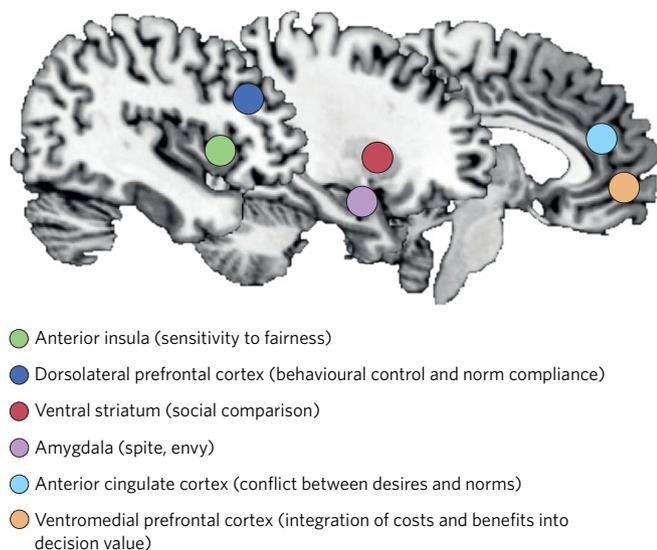
- 🟢 Anterior insula (sensitivity to fairness)
- 🔵 Dorsolateral prefrontal cortex (behavioural control and norm compliance)
- 🔴 Ventral striatum (social comparison)
- 🟣 Amygdala (spite, envy)
- 🔵 Anterior cingulate cortex (conflict between desires and norms)
- 🟠 Ventromedial prefrontal cortex (integration of costs and benefits into decision value)

Figure 3 | Brain regions supporting fairness behaviour in children and adults.

Additionally, the subjective experience of envy has been shown to relate to disadvantageous inequity aversion in children as young as 6 years[59]. Other work suggests that disadvantageous inequity aversion in children is motivated by spite, since children only reject a small reward when doing so will reduce the partner's larger reward[33]. Having less than others thus elicits a negative emotional state, which drives the response to reject such divisions. This links well with findings in adults, where activity of the amygdala, a brain region implicated in various emotional states[60], particularly negative ones[61,62], correlated positively with subjects' willingness to reduce others' outcomes to create equality[63]. The experience of negative emotions is thus a key component for the emergence of disadvantageous inequity aversion and second-party punishment, the earliest developmental signatures of fairness. Importantly, however, these aspects of fairness are still quite limited, as they focus on fairness towards oneself.

The role of conflict resolution, behavioural control and social norm compliance in creating fairness for oneself and others. Creating fairness entails incurring a cost, which in turn creates a conflict with self-interest. Evidence of such a conflict can be detected in children's reaction times. For instance, in the Inequity Game (see Box 2), children of all ages take longer to decide when faced with disadvantageous inequity than equity.

Resolving the conflict between self-interest and fairness can be achieved by complying with social norms, which requires behavioural control. Such conflict is evident in all four signatures of human fairness. However, the neural mechanisms of conflict resolution and, more importantly, behavioural control and social norm compliance have been most extensively investigated in the context of punishing others' fairness violations. Several studies in adults report a consistently activated system of brain regions implicated in resolving conflicts during both second-party[57,64–66] and third-party punishment tasks[56,67–69]. In addition to the anterior insula, this network includes the ACC and the DLPFC. Activity of the ACC is considered to reflect conflict detection and resolution[70] and, in the context of punishing fairness violations, is argued to reflect the conflict between self-interest and restoring fairness[65].

Both fMRI studies and studies using non-invasive brain stimulation such as transcranial magnetic stimulation (TMS) and direct current stimulation (tDCS) suggest that activity of the DLPFC and VMPFC is crucial for social norm compliance[65,71–75]. These findings suggest that DLPFC is critical in implementing norm-related

behaviour by means of exerting a modulating effect on VMPFC, a brain region implicated in integrating costs and benefits of multiple decision options[73]. Such an interpretation of DLPFC is supported by studies of the neurocognitive underpinnings of other types of norm violations (that is, criminal acts) in a third-party context, which also suggest that DLPFC is involved in norm enforcement[76,77].

The development of behavioural control has been shown to explain the emergence of punishing norm violations. For instance, in one study, 6-year-olds claimed that they would reject an unfair offer when asked hypothetically, yet when faced with real situations they accepted such offers readily[44]. This stands in contrast to older children who do as they say more often. Such a developmental dissociation between judgments of fairness and actual behaviour has been found in other contexts as well. Most notably, in Dictator Games children as young as 3 years recognize an equal split as a norm but keep more for themselves, suggesting that young children do not view fairness norms as obligatory[26]. By about 6–8 years of age, children are more likely to follow the stated norm in this task, but compliance with the norm appears to be linked to behavioural control at least until 12 years of age[78]. Finally, a recent study provides direct evidence for the role of developmental changes in DLPFC and associated behavioural control in accounting for changes in social norm compliance. Proposer offers in the Ultimatum Game compared with the Dictator Game in children aged 7 to 13 years were strongly correlated with behavioural control, and in turn supported by both the function and structure of the left DLPFC[44]. Behavioural control is thus likely to be an important mechanism that supports the emergence of norm compliance required for punishment of norm violations.

Behavioural control is also required when rejecting advantageous distributions, given that a relatively large sum is sacrificed for the sake of equality. Notably, once advantageous inequity aversion emerges around 8 years of age, children take longer to make their decisions, suggesting a conflict between self-interest and norm compliance[28]. Sacrificing an advantage and resolving this conflict in favour of norm compliance depends on an improvement or shift in the quality of behavioural control. One such developmental shift is that behavioural control becomes increasingly more proactive, flexible and internalized[53]. This leads to the hypothesis that as behavioural control increases, fairness norms will be followed in a wider set of contexts and also applied not just to others but also to oneself[79].

The role of valuing equal outcomes in creating fairness for others. Finally, it is necessary to consider the mechanisms that allow children to incur costs on behalf of another, as is the case both in third-party punishment and in rejections of advantageous distributions. In addition to an increased ability to comply with norms, we argue that this shift can be partially explained by an increasing value placed on fair outcomes *per se*. This is evidenced in some countries by children around 8 years of age who show clear signs of advantageous inequity aversion and also explicitly claim that they do so because it did not seem fair to have more than someone else[28]. The anterior insula has long been noted to be particularly sensitive to fairness violations[64]. It has also been implicated in responding to the emotional experiences of others[80] and thus constitutes a region highly suited to respond to others falling victim to fairness violations[81,82]. It is therefore noteworthy that, in adults, both in the context of third-party punishment and advantageous inequity aversion, the anterior insula plays a critical role, particularly with regards to its connectivity to other brain regions, such as the striatum and the VMPFC[56,63,69]. We propose that studying the developmental changes in the functional recruitment and integration of the anterior insula into the neural circuitry underlying fairness behaviours is a promising avenue to explain the emergence and development of fairness in childhood.
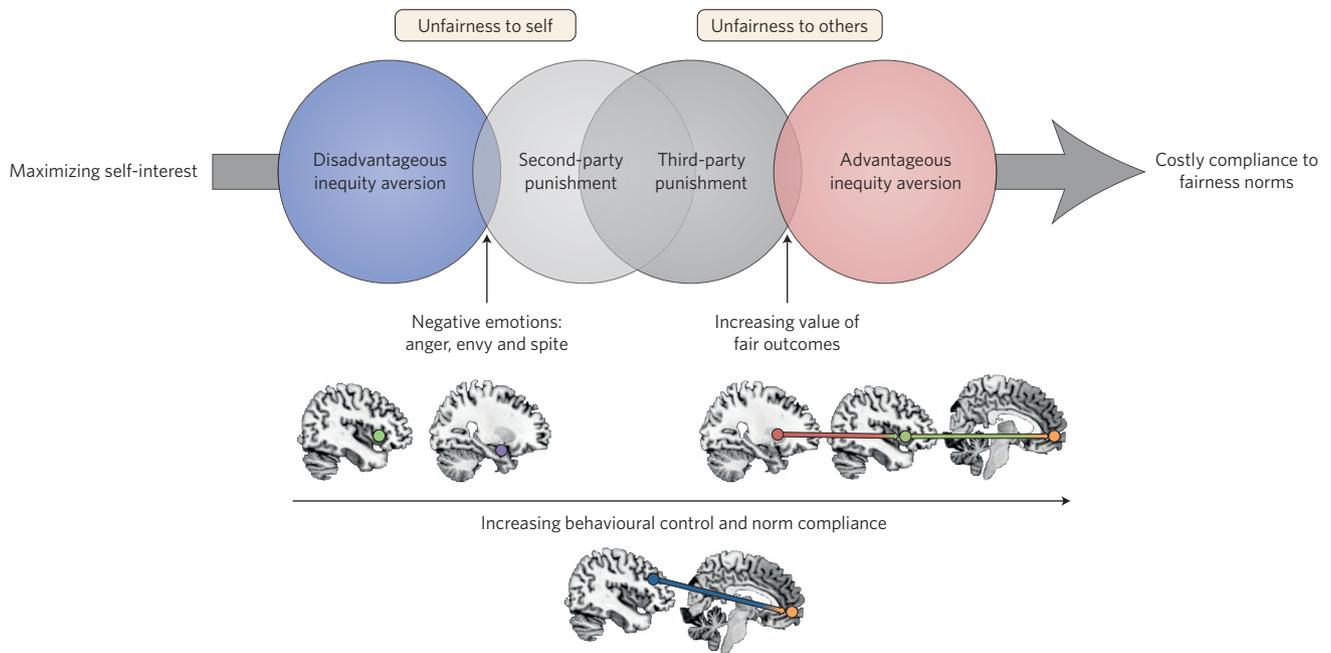
Figure 4 | Diagram showing the developmental changes in the emergence of fairness behaviours in childhood along with the neurocognitive mechanisms hypothesized to explain the observed shifts from fairness for oneself to fairness for others. Negative emotions such as anger, envy and spite drive reactions towards unfairness to self, presumably mediated by the amygdala (purple) and the anterior insula (green). Incurring costs to follow and enforce social norms and achieve fair outcomes is related to the development of behavioural control, mediated by the development of connectivity between DLPFC (blue) and VMPFC (orange). Finally, we hypothesize that an increased value placed on fairness drives reactions towards unfairness to others, which is linked to increased connectivity of the anterior insula with the striatum (red) and, more importantly, VMPFC.

Synthesis of the mechanisms supporting the emergence of fairness behaviours in children. In sum, a set of circumscribed mechanisms supports the emergence of distinct signatures of human fairness (Fig. 4). When taken together with behavioural evidence, neural data begin to paint a more complete picture of the emergence of fairness in childhood and provide critical insight into the mechanisms that support this process. Strikingly, the dissociation between disadvantageous and advantageous inequity aversion observed across development is mirrored by unique supporting mechanisms. To synthesize what we know so far: first, the appearance of disadvantageous inequity aversion is driven by self-centered motives of not wanting to have less than others. Second, punishment of unfairness in others appears to occur as a function of increased behavioural control, which supports norm compliance. Finally, since incurring costs to forgo an unfair advantage and punishing on behalf of others seem to rely on a heightened sense of fairness, we propose that both behaviours are likely to be coded, at least in part, in the insula. Based on adult data, connectivity between the insula, striatum and VMPFC appears to be a critical neural mechanism for both advantageous inequity aversion and third-party punishment. Connectivity between distinct neural systems (Box 3) is a protracted developmental process. This may help explain why advantageous inequity aversion, which we hypothesize to be reliant on this mechanism, emerges later in development. However, before putting too much weight on this claim, it is important to test this hypothesis with developmental samples.

Conclusions and future directions

The evidence reviewed here offers an exciting look at the emergence of fairness behaviour in human ontogeny and the mechanisms supporting it. However, our working definition of fairness is admittedly a lean one: namely, costly responses to unequal resource distributions. Of course, fairness judgments and behaviours are not restricted to these cases and an important direction for future work

is to understand how children make decisions about other forms of fairness such as procedural[83] and retributive justice[84]. Additionally, we know that a suite of factors influences fairness behaviour in adults and children, including reputational concerns[85], perspective taking[86], in-group favouritism[50], as well as cues of deservedness such as merit[87] and need[7]. Integrating evidence about the extent to which these factors affect fairness behaviour with burgeoning neuroscientific data on these same aspects of human cognition (for example, reputation[88], perspective taking[89] and parochialism[90]) is an exciting avenue for future work. Moreover, we know little about the extent to which inter-individual variation in different aspects of children's developing cognition (for example, empathy, numerical knowledge and working memory) influences children's fairness behaviour. Characterizing influences on fairness at the individual level will provide detailed insight into the cognitive mechanisms that support its development.

The human sense of fairness has its roots in ontogeny, yet shows striking developmental changes with age. In early childhood, children avoid having less than others (disadvantageous inequity aversion), by mid-childhood they punish unfairness in both second- and third-party contexts and by late childhood they avoid having more than others (advantageous inequity aversion). Recent neuroscientific evidence offers possible explanations for this pattern of development. Specifically, improvements in behavioural control and a growing value placed on fair outcomes may help explain how children shift from a seemingly myopic focus on ensuring fairness for self towards a more generalized sense of fairness which entails the costly enforcement of and compliance to fairness norms. More broadly, we show that the integration of evidence from developmental psychology and cognitive neuroscience can offer fundamental insights into the mechanisms supporting the human sense of fairness.

References

1. Tinbergen, N. On aims and methods of ethology. *Z. Tierpsychol.* **20,** 410–433 (1963).
2. Bateson, P. & Laland, K. N. Tinbergen's four questions: an appreciation and an update. *Trends Ecol. Evol.* **28,** 712–718 (2013).
3. West-Eberhard, M. J. *Developmental Plasticity and Evolution* (Oxford Univ. Press, 2003).
4. Aristotle. *The Nicomachean Ethics* (Harvard Univ. Press, 1999).
5. Rawls, J. *A Theory of Justice* (Harvard Univ. Press, 1971).
6. Adams, J. S. Inequity in social exchange. *Adv. Exp. Soc. Psychol.* **2,** 267–299 (1965).
7. Deutsch, M. Equity, equality, and need: what determines which value will be used as the basis of distributive justice? *J. Soc. Issues* **31,** 137–149 (1975).
8. Loewenstein, G. F., Thompson, L. & Bazerman, M. H. Social utility and decision making in interpersonal contexts. *J. Pers. Soc. Psychol.* **57,** 426–441 (1989).
9. Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R. & Smirnov, O. Egalitarian motives in humans. *Nature* **446,** 794–796 (2007).
10. Camerer, C. *Behavioral Game Theory: Experiments in Strategic Interaction* (Princeton Univ. Press, 2003).
11. Fehr, E. & Schmidt, K. M. A theory of fairness, competition, and cooperation. *Q. J. Econ.* **114,** 817–868 (1999).
12. Güth, W., Schmittberger, R. & Schwarze, B. An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* **3,** 367–388 (1982).
13. Henrich, J. *et al.* "Economic man" in cross-cultural perspective: behavioral experiments in 15 small-scale societies. *Behav. Brain Sci.* **28,** 795–815 (2005).
14. Fehr, E. & Fischbacher, U. Third-party punishment and social norms. *Evol. Hum. Behav.* **25,** 63–87 (2004).
15. Henrich, J. *et al.* Costly punishment across human societies. *Science* **312,** 1767–1770 (2006).
16. Brosnan, S. F. & de Waal, F. B. M. Evolution of responses to (un)fairness. *Science* **346,** 1251776 (2014).
17. Bräuer, J. & Hanus, D. Fairness in non-human primates? *Soc. Just. Res.* **25,** 256–276 (2012).
18. McAuliffe, K. & Santos, L. R. in *The Atlas of Moral Psychology: Mapping Good and Evil in the Mind* (eds Gray, K. & Graham, J.) (Guilford Press, in the press).
19. Jensen, K., Call, J. & Tomasello, M. Chimpanzees are rational maximizers in an ultimatum game. *Science* **318,** 107–109 (2007).
20. Proctor, D., Williamson, R. A., de Waal, F. B. M. & Brosnan, S. F. Chimpanzees play the ultimatum game. *Proc. Natl Acad. Sci. USA* **110,** 2070–2075 (2013).
21. Riedl, K., Jensen, K., Call, J. & Tomasello, M. No third-party punishment in chimpanzees. *Proc. Natl Acad. Sci. USA* **109,** 14824–14829 (2012).
22. Fehr, E., & Fischbacher, U. Social norms and human cooperation. *Trends Cogn. Sci.* **8,** 185–190 (2004).
23. Geraci, A. & Surian, L. The developmental roots of fairness: infants' reactions to equal and unequal distributions of resources. *Dev. Sci.* **14,** 1012–1020 (2011).
24. Schmidt, M. F. H. & Sommerville, J. A. Fairness expectations and altruistic sharing in 15-month-old human infants. *PLoS ONE* **6,** 1–7 (2011).
25. Sloane, S., Baillargeon, R. & Premack, D. Do infants have a sense of fairness? *Psychol. Sci.* **23,** 196–204 (2012).
26. Smith, C. E., Blake, P. R. & Harris, P. L. I should but I won't: why young children endorse norms of fair sharing but do not follow them. *PLoS ONE* **8,** e59510 (2013).
27. LoBue, V., Nishida, T., Chiong, C., DeLoache, J. S. & Haidt, J. When getting something good is bad: even three-year-olds react to inequality. *Soc. Dev.* **20,** 154–170 (2010).
28. Blake, P. R. & McAuliffe, K. I had so much it didn't seem fair: eight-year-olds reject two forms of inequity. *Cognition* **120,** 215–224 (2011).
29. Fehr, E., Bernhard, H. & Rockenbach, B. Egalitarianism in young children. *Nature* **454,** 1079–1083 (2008).
30. Shaw, A., Choshen-Hillel, S., & Caruso, E. The development of inequity aversion: understanding when (and why) people give others the bigger piece of the pie. *Psychol. Sci.* **27,** 1352–1359 (2016).
31. Chernyak, N., & Kushnir, T. Giving preschoolers choice increases sharing behavior. *Psychol. Sci.* **24,** 1971–1979 (2013).
32. McAuliffe, K., Blake, P. R., Kim, G., Wrangham, R. W. & Warneken, F. Social influences on inequity aversion in children. *PLoS ONE* **8,** e80966 (2013).
33. McAuliffe, K., Blake, P. R. & Warneken, F. Children reject inequity out of spite. *Biol. Lett.* **10,** 20140743 (2014).
34. Blake, P. R., McAuliffe, K., Corbit, J., Callaghan, T. C. & Barry, O. The ontogeny of fairness in seven societies. *Nature* **528,** 258–261 (2015).
35. Wittig, M., Jensen, K. & Tomasello, M. Five-year-olds understand fair as equal in a mini-ultimatum game. *J. Exp. Child Psychol.* **116,** 324–337 (2013).
36. Lucas, M. M., Wagner, L. & Chow, C. Fair game: the intuitive economics of resource exchange in four-year olds. *J. Soc. Evol. Cult. Psychol.* **2,** 74–88 (2008).
37. Bereby-Meyer, Y. & Fiks, S. Changes in negative reciprocity as a function of age. *J. Behav. Decis. Making* **26,** 397–403 (2013).
38. Gummerum, M. & Chu, M. T. Outcomes and intentions in children's, adolescents', and adults' second- and third-party punishment behavior. *Cognition* **133,** 97–103 (2014).
39. Sutter, M. Outcomes versus intentions: on the nature of fair behavior and its development with age. *J. Econ. Psychol.* **28,** 69–78 (2007).
40. Sally, D. & Hill, E. The development of interpersonal strategy: autism, theory-of-mind, cooperation and fairness. *J. Econ. Psychol.* **27,** 73–97 (2006).
41. Castelli, I., Massaro, D., Sanfey, A. G. & Marchetti, A. Fairness and intentionality in children's decision-making. *Int. Rev. Econ.* **57,** 269–288 (2010).
42. Castelli, I., Massaro, D., Bicchieri, C., Chavez, A. & Marchetti, A. Fairness norms and theory of mind in an ultimatum game: judgments, offers, and decisions in school-aged children. *PLoS ONE* **9,** e105024 (2014).
43. Harbaugh, W., Krause, K. & Liday, S. J. Bargaining by children. Preprint at http://dx.doi.org/10.2139/ssrn.436504 (2003).
44. Steinbeis, N., Bernhardt, B. C. & Singer, T. Impulse control and underlying functions of the left DLPFC mediate age-related and age-independent individual differences in strategic social behavior. *Neuron* **73,** 1040–1051 (2012).
45. Güroğlu, B., van den Bos, W. & Crone, E. A. Fairness considerations: increasing understanding of intentionality during adolescence. *J. Exp. Child Psychol.* **104,** 398–409 (2009).
46. Hamlin, J. K. & Wynn, K. Young infants prefer prosocial to antisocial others. *Cogn. Dev.* **26,** 30–39 (2011).
47. Hamlin, J. K., Wynn, K., Bloom, P. & Mahajan, N. How infants and toddlers react to antisocial others. *Proc. Natl Acad. Sci. USA* **108,** 19931–19936 (2011).
48. McAuliffe, K., Jordan, J. J. & Warneken, F. Costly third-party punishment in young children. *Cognition* **134,** 1–10 (2015).
49. Salali, G. D., Juda, M. & Henrich, J. Transmission and development of costly punishment in children. *Evol. Hum. Behav.* **36,** 86–94 (2015).
50. Jordan, J. J., McAuliffe, K. & Warneken, F. Development of in-group favoritism in children's third-party punishment of selfishness. *Proc. Natl Acad. Sci. USA* **111,** 12710–12715 (2014).
51. Sheskin, M., Bloom, P. & Wynn, K. Anti-equality: social comparison in young children. *Cognition* **130,** 152–156 (2014).
52. Shaw, A. & Olson, K. R. Children discard a resource to avoid inequity. *J. Exp. Psychol. Gen.* **141,** 382–395 (2012).
53. Munakata, Y., Snyder, H. R. & Chatham, C. H. Developing cognitive control: three key transitions. *Curr. Dir. Psychol. Sci.* **21,** 71–77 (2012).
54. Zelazo, P. D. *et al.* II. NIH Toolbox Cognition Battery (CB): measuring executive function and attention. *Monogr. Soc. Res. Child Dev.* **78,** 16–33 (2013).
55. Fehr, E. & Camerer, C. F. Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn. Sci.* **11,** 419–427 (2007).
56. Feng, C. *et al.* Diffusion of responsibility attenuates altruistic punishment: a functional magnetic resonance imaging effective connectivity study. *Hum. Brain Mapp.* **37,** 663–677 (2016).
57. Gabay, A. S., Radua, J., Kempton, M. J. & Mehta, M. A. The Ultimatum Game and the brain: a meta-analysis of neuroimaging studies. *Neurosci. Biobehav. Rev.* **47,** 549–558 (2014).
58. Ruff, C. C. & Fehr, E. The neurobiology of rewards and values in social decision making. *Nat. Rev. Neurosci.* **15,** 549–562 (2014).
59. Steinbeis, N. & Singer, T. The effects of social comparison on social emotions and behavior during childhood: the ontogeny of envy and Schadenfreude predicts developmental changes in equity-related decisions. *J. Exp. Child Psychol.* **115,** 198–209 (2013).
60. Phelps, E. A. Emotion and cognition: insights from studies of the human amygdala. *Annu. Rev. Psychol.* **57,** 27–53 (2006).
61. Barrett, L. F., Bliss-Moreau, E., Duncan, S. L., Rauch, S. L. & Wright, C. I. The amygdala and the experience of affect. *Soc. Cogn. Affect. Neurosci.* **2,** 73–83 (2007).
62. Feinstein, J. S., Adolphs, R., Damasio, A. & Tranel, D. The human amygdala and the induction and experience of fear. *Curr. Biol.* **21,** 34–38 (2011).
63. Yu, R., Calder, A. J. & Mobbs, D. Overlapping and distinct representations of advantageous and disadvantageous inequality. *Hum. Brain Mapp.* **35,** 3290–3301 (2014).
64. Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E. & Cohen, J. D. The neural basis of economic decision-making in the Ultimatum Game. *Science* **300,** 1755–1758 (2003).
65. Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C. & Fehr, E. Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nat. Neurosci.* **14,** 1468–1474 (2011).
66. Güroğlu, B., van den Bos, W., Rombouts, S. A. R. B. & Crone, E. A. Unfair? It depends: neural correlates of fairness in social context. *Soc. Cogn. Affect. Neurosci.* **5,** 414–423 (2010).
67. Corradi-Dell'Acqua, C., Civai, C., Rumiati, R. I. & Fink, G. R. Disentangling self- and fairness-related neural mechanisms involved in the Ultimatum Game: an fMRI study. *Soc. Cogn. Affect. Neurosci.* **8,** 424–431 (2013).

68. Civai, C., Crescentini, C., Rustichini, A. & Rumiati, R. I. Equality versus self-interest in the brain: differential roles of anterior insula and medial prefrontal cortex. *NeuroImage* **62,** 102–112 (2012).

69. Zhong, S., Chark, R., Hsu, M. & Chew, S. H. Computational substrates of social norm enforcement by unaffected third parties. *NeuroImage* **129,** 95–104 (2016).

70. Carter, C., Van Veen, V., Botvinick, M., Cohen, J. & Stenger, V. A. Conflict and the evaluative functions of the anterior cingulate: converging evidence from event-related fMRI and high density ERP. *NeuroImage* **13,** S305–S305 (2001).

71. Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V. & Fehr, E. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* **314,** 829–832 (2006).

72. van't Wout, M., Kahn, R. S., Sanfey, A. G. & Aleman, A. Repetitive transcranial magnetic stimulation over the right dorsolateral prefrontal cortex affects strategic decision-making. *Neuroreport* **16,** 1849–1852 (2005).

73. de Quervain, D. *et al.* The neural basis of altruistic punishment. *Science* **305,** 1254–1258 (2004).

74. Ruff, C. C., Ugazio, G. & Fehr, E. Changing social norm compliance with noninvasive brain stimulation. *Science* **342,** 482–484 (2013).

75. Spitzer, M., Fischbacher, U., Herrnberger, B., Grön, G. & Fehr, E. The neural signature of social norm compliance. *Neuron* **56,** 185–196 (2007).

76. Buckholtz, J. W. & Marois, R. The roots of modern justice: cognitive and neural foundations of social norms and their enforcement. *Nat. Neurosci.* **15,** 655–661 (2012).

77. Buckholtz, J. W. *et al.* From blame to punishment: disrupting prefrontal cortex activity reveals norm enforcement mechanisms. *Neuron* **87,** 1369–1380 (2015).

78. Blake, P. R., Piovesan, M., Montinari, N. & Warneken, F. Prosocial norms in the classroom: the role of self-regulation in following norms of giving. *J. Econ. Behav. Organ.* **115,** 18–29 (2015).

79. Steinbeis, N. & Crone, E. A. The link between cognitive control and decision-making across child and adolescent development. *Curr. Opin. Behav. Sci.* **10,** 28–32 (2016).

80. Lamm, C., Decety, J. & Singer, T. Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage* **54,** 2492–2502 (2011).

81. Dawes, C. T. *et al.* Neural basis of egalitarian behavior. *Proc. Natl Acad. Sci. USA* **109,** 6479–6483 (2012).

82. Corradi-Dell'Acqua, C., Tusche, A., Vuilleumier, P. & Singer, T. Cross-modal representations of first-hand and vicarious pain, disgust and fairness in insular and cingulate cortex. *Nat. Commun.* **7,** 10904 (2016).

83. Grocke, P., Rossano, F. & Tomasello, M. Procedural justice in children: preschoolers accept unequal resource distributions if the procedure provides equal opportunities. *J. Exp. Child Psychol.* **140,** 197–210 (2015).

84. Smith, C. E. & Warneken, F. Children's reasoning about distributive and retributive justice across development. *Dev. Psychol.* **52,** 613–628 (2016).

85. Shaw, A. *et al.* Children develop a veil of fairness. *J. Exp. Psychol. Gen.* **143,** 363–375 (2014).

86. Takagishi, H., Kameshima, S., Schug, J., Koizumi, M. & Yamagishi, T. Theory of mind enhances preference for fairness. *J. Exp. Child Psychol.* **105,** 130–137 (2010).

87. Kanngiesser, P. & Warneken, F. Young children consider merit when sharing resources with others. *PLoS ONE* **7,** e43979 (2012).

88. Knoch, D. *et al.* Disrupting the prefrontal cortex diminishes the human ability to build a good reputation. *Proc. Natl Acad. Sci. USA* **106,** 20895–20899 (2009).

89. Saxe, R. & Kanwisher, N. People thinking about thinking people. The role of the temporo-parietal junction in 'theory of mind'. *NeuroImage* **19,** 1835–1842 (2003).

90. Baumgartner, T., Schiller, B., Rieskamp, J., Gianotti, L. R. R. & Knoch, D. Diminishing parochialism in intergroup conflict by disrupting the right temporo-parietal junction. *Soc. Cogn. Affect. Neurosci.* **9,** 653–660 (2014).

91. Thompson, C., Barresi, J. & Moore, C. The development of future-oriented prudence and altruism in preschoolers. *Cogn. Dev.* **12,** 199–212 (1997).

92. Gogtay, N. *et al.* Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl Acad. Sci. USA* **101,** 8174–8179 (2004).

93. Shaw, P. *et al.* Intellectual ability and cortical development in children and adolescents. *Nature* **440,** 676–679 (2006).

94. Sowell, E. R. Longitudinal mapping of cortical thickness and brain growth in normal children. *J. Neurosci.* **24,** 8223–8231 (2004).

95. Lebel, C. & Beaulieu, C. Longitudinal development of human brain wiring continues from childhood into adulthood. *J. Neurosci.* **31,** 10937–10947 (2011).

96. Power, J. D., Fair, D. A., Schlaggar, B. L. & Petersen, S. E. The development of human functional brain networks. *Neuron* **67,** 735–748 (2010).

97. Zielinski, B. A., Gennatas, E. D., Zhou, J. & Seeley W. W. Network-level structural covariance in the developing brain. *Proc. Natl Acad. Sci. USA* **107,** 18191–18196 (2010).

98. Kelly, A. M. C. *et al.* Development of anterior cingulate functional connectivity from late childhood to early adulthood. *Cereb. Cortex* **19,** 640–657 (2009).

99. Fair, D. A. *et al.* Development of distinct control networks through segregation and integration. *Proc. Natl Acad. Sci. USA* **104,** 13507–13512 (2007).

100. Supekar, K., Musen, M. & Menon, V. Development of large-scale functional brain networks in children. *PLoS Biol.* **7,** e1000157 (2009).

101. Uddin, L. Q., Supekar, K. S., Ryali, S. & Menon, V. Dynamic reconfiguration of structural and functional connectivity across core neurocognitive brain networks with development. *J. Neurosci.* **31,** 18578–18589 (2011).

102. Luna, B., Marek, S., Larsen, B., Tervo-Clemmens, B. & Chahal, R. An integrative model of the maturation of cognitive control. *Annu. Rev. Neurosci.* **38,** 151–170 (2015).

103. Marek, S., Hwang, K., Foran, W., Hallquist, M. N. & Luna, B. The contribution of network organization and integration to the development of cognitive control. *PLoS Biol.* **13,** e1002328 (2015).

Acknowledgements

Additional information

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence should be addressed to K.M.

How to cite this article: McAuliffe, K., Blake, P. R., Steinbeis, N. & Warneken, F. The developmental foundations of human fairness. *Nat. Hum. Behav.* **1,** 0042 (2017).

Competing interests

The authors declare no competing interests.