Outcomes versus intentions in fairness-related decision making: School-aged children's decisions are just like those of adults

Published in:
Journal of Experimental Child Psychology

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Outcomes Versus Intentions in Fairness-Related Decision Making:

School-Aged Children's Decisions Are Just Like Those of Adults

Agnieszka J. Jaroslawska[a], Teresa McCormack[a], Patrick Burns[a], and Eugene M. Caruso[b]

[a]Queen's University Belfast, [b]University of California Los Angeles

Abstract

The notion of what constitutes fairness has been assumed to change during childhood, in line with a marked shift from outcome-based to intention-based moral reasoning. However, the precise developmental profile of such a shift is still subject to debate. This study sought to determine the age at which the perceived intentions of others begin to influence fairness-related decision-making in children (aged 6-8 and 9-11) and adolescents (aged 14-15) in the context of the mini-ultimatum game. The mini-ultimatum game has a forced-choice design whereby a proposer has to select one of two pre-determined offers that a responder can either accept or reject. Due to these constraints the procedure measures sensitivity to unfair intentions in addition to unfair outcomes. Participants had to make judgments about how likely they would be to reject various offers, how fair they judged these offers to be, and the emotion they experienced when thinking about the offer. Contrary to previous published reports, we found that even 6- to 8-year-olds employed a sophisticated notion of fairness that took into account the alternatives the proposer had available to them. Crucially, decision-making did not differ as a function of age. A further, and novel, aim was to trace the developmental origins of temporal asymmetries in judgments about fairness by testing the implications of adopting a past or a future temporal perspective. Across all ages, we found no evidence that fairness-based decision-making varies as a function of temporal location.

Keywords: decision-making, development, fairness, intentionality, temporal asymmetry

Outcomes Versus Intentions in Fairness-Related Decision Making:

School-Aged Children's Decisions Are Just Like Those of Adults

Young children know a great deal about fairness. Recent studies tracing the development of moral reasoning have demonstrated that, under appropriate circumstances, both toddlers and pre-verbal infants are sensitive to the basic principles of equal distribution of resources (see Blake, McAuliffe, & Warneken, 2014, for a review), with some studies suggesting that infants between 9 and 21 months expect equality (e.g., in Dawkins, Sloane, & Baillargeon, 2019; in Sloane, Baillargeon, & Premack, 2012; in Sommerville, Schmidt, Yun, & Burns, 2013) and prefer individuals who enact equal allocations (Geraci & Surian, 2011; Margoni & Surian, 2018). Preschoolers also notice and respond with negative emotional expression to unfair distributions made by another person (LoBue, Nishida, Chiong, DeLoache, & Haidt, 2011). These findings indicate that a basic affective response to unequal distributions is in place early in development.

However, a basic aversion to inequity can be viewed as just the first developmental step in the emergence of a fully-fledged notion of fairness. Some models of fairness in social decision making assume there are other important developmental changes, such that fairness is conceptualized initially as inequity aversion and based solely on social comparisons between one's own and others' outcomes (e.g., Radke, Güroğlu, & de Bruijn, 2012; see also Gummerum & Chu, 2014). It is only later in development that children begin to incorporate contextual information and intentionality into their assessments of fairness in social decision making. In the case of resource allocation decisions, Radke et al. (2012) use the term contextual information to mean information about what alternative options were available to the individual who made a resource allocation decision, other than the option that they actually selected. Incorporating intentionality into fairness considerations involves evaluating the extent to which the decider deliberately intended to act fairly or unfairly. In resource

allocation tasks, intentionality judgments are closely linked to context: For example, whether a decider had other, more equitable, options available will influence whether an inequitable division of resources is deemed to be intentional (Falk, Fehr, & Fischbacher, 2003; Falk & Fischbacher, 2006). Radke et al. (2012) argue that considering contextual and intentional information requires advanced cognitive competencies over and above those required for a basic aversion to inequity.

More specifically, Radke et al. (2012) claim that considering contextual information involves thinking counterfactually about whether alternative options available to the decider would have yielded better or worse outcomes, and considering intentions requires perspective-taking skills. By perspective-taking skills, what Radke et al. have in mind is specifically the ability to actively keep track during social interactions of the psychological motives that drive other people's behaviours, including monitoring whether an agent is aware of the consequences of their actions. Both of these types of abilities – counterfactual thinking and perspective-taking – are thought to emerge relatively early in basic forms but to continue to improve across middle to late childhood. Although there is some evidence that three-year-olds can make simple counterfactual judgments (Harris, German, & Mills, 1996), several theorists argue that genuine counterfactual reasoning emerges later in development (for review Beck & Riggs, 2014; Rafetseder, Cristi-Vargas, & Perner, 2010), and indeed that children may not make use of counterfactual information to evaluate an outcome until around 6 years of age (O'Connor, McCormack, & Feeney, 2014) or even later (Rafetseder, Schwitalla, & Perner, 2013). Similarly, although the ability to understand, represent, and refer to, mental states of others is assumed to develop by 4-to-5 years of age (Wellman, Cross, & Watson, 2001), perspective-taking skills, including the ability to spontaneously track other's mental states during social interactions, improve gradually over the course of childhood and early adolescence (see Crone, 2013 for review; Dumontheil, Apperly, & Blakemore, 2010;

Valle, Massaro, Castelli, & Marchetti, 2015). Given this, more sophisticated fairness considerations may emerge developmentally, as the ability to reason counterfactually and engage in perspective-taking about intentions improves (cf. Castelli, Massaro, Bicchieri, Chavez, & Marchetti, 2014a). As discussed below, though, the development profile of changes in the nature of fairness considerations has not been clearly established.

Research with adults suggests that there is a further, and perhaps surprising, type of contextual factor that affects judgments of fairness: the temporal context in which such judgments are elicited. There is a growing body of evidence demonstrating that adults show a number of past-future asymmetries in their judgments that suggests that they are biased toward the future. For instance, compared to identical events in the past, events in the future are more highly valued (e.g., Caruso, Gilbert, & Wilson, 2008) and evoke more intense emotion (Van Boven & Ashworth, 2007). Moral evaluations, including fairness judgments, also seem to be affected by temporal context: For example, Caruso (2010) found that adults were more likely to judge an inequitable distribution of resources as unfair if it was due to take place in the future rather than the past, which he argues is linked to the stronger emotions aroused when considering future events. Some recent research suggests that children, like adults, also feel stronger emotions when considering the future (Burns, McCormack, Jaroslawska, Fitzpatrick, McGourty, & Caruso, 2019). However, no studies to date have examined whether children's fairness judgments are affected by temporal context.

The aims of the current study are twofold. First, it aimed to trace the developmental shift from outcome-based to intention-based reasoning about fairness using the mini-ultimatum game and to determine the age at which the perceived intentions of others begin to influence judgments of fairness in children and adolescents. Second, our aim was to explore, for the first time, the developmental origins of temporal asymmetries in children's and adolescents' judgments about fairness. We outline the rationale for these manipulations in

OUTCOMES VERSUS INTENTIONS IN FAIRNESS-RELATED DECISION MAKING

more detail below, but first we present a brief review of previous research motivating the current study.

Measuring Fairness Considerations

The ways in which social decision making is guided by considerations of fairness has been most extensively studied in adults using versions of the ultimatum game (Güth & Kocher, 2014; Güth, Schmittberger, & Schwarze, 1982). In this game two players are instructed to divide a sum of money. The first player (i.e., the proposer) decides how the money is distributed; the second player (i.e., the responder) decides whether to accept or reject the suggested split. If the responder accepts the offer, the players receive the proposed amounts. If the responder rejects the offer, both players receive nothing. The standard economic model dictates that, to maximize their payoffs, the proposer should offer the smallest amount of money, and the responder should accept all offers, no matter how unfair. Ultimately, a small sum of money is better than no money. However, contrary to this prediction, a proportion of unfair offers are reliably rejected, indicating that individuals would rather suffer a financial loss than accept an unequal division of resources (see Camerer, 2003 for a review; Marchetti et al., 2019; Sanfey Rilling, Aronson, Nystrom, & Cohen, 2003; Van Dijk & Vermunt, 2000).

Two influential theories offer an explanation for this apparently irrational behaviour. The theory of negative reciprocity (e.g., Falk & Fischbacher, 2006; Rabin, 1993) focuses on intentions and describes rejections as a tool to punish the unfair proposer. Rather than maximizing self-interest, the responder is punishing socially inappropriate behaviour for the greater good of the group (Fehr & Fischbacher, 2003). The theory of inequality aversion (e.g., Fehr & Schmidt, 1999; see also Bicchieri & Chavez, 2010) claims that people are naturally averse to unequal distributions, especially when disadvantageous. Inequality aversion is thought to reflect an emotional reaction to unfairness that overrides purely

economic decision processes. Indeed, being treated unfairly leads to anger (Pillutla &

Murninghan, 1996) and has been associated with activation of neural networks involved in

conflict and negative affect (Güroğlu, van den Bos, Rombouts, & Crone, 2010; Güroğlu, van

den Bos, van Dijk, Rombouts, & Crone, 2011). Brain imaging studies confirm that areas

associated with emotion show heightened activity when people reject unfair offers (Sanfey et

al., 2003) and that tolerating unfair treatment requires the suppression of negative emotions

(Tabibnia, Satpute, & Lieberman, 2008).

The ultimatum game has been used as a tool for examining the nature of fairness

considerations in decision-making. In the context of the ultimatum game, a distinction has

been made between outcome-based fairness assessments (i.e., comparing one's own gains

and losses relative to those of others) and intention-based fairness assessments (i.e., taking

into account the motives lying behind the outcomes). For example, a responder may be more

willing to accept an unfair offer from a computer (i.e., determined by a random number

generator) than from another player (e.g., Blount, 1995). A modified version of the ultimatum

game, the so-called mini-ultimatum game, probes the context of unfair offers. Pairing a

disadvantageous offer with different alternatives allows researchers to investigate how

responders react to unfair offers depending on the perceived intention with which they are

given (Falk et al., 2003; Falk & Fischbacher, 2006). Consequently, in the mini-ultimatum

game, players are confronted with situations in which the simple rule of thumb that fair

means equal does not always apply. Because of its forced-choice design, the mini-ultimatum

game measures sensitivity to unfair intentions in addition to unfair outcomes.

In the original mini-ultimatum paradigm designed by Falk et al. (2003), two

allocations of points are presented as a forced choice. Across four trials, one of the outcomes

is always disadvantageous (i.e., 8 points for the proposer and 2 points for the responder). The

alternative offers are: an equal outcome (i.e., 5 points for both players), a generous outcome

(i.e., 2 points for the proposer and 8 points for the responder), an equally unfair outcome (i.e.,

8 points for the proposer and 2 points for the responder), and a hyper-unfair outcome (10

points for the proposer and 0 points for the responder). Importantly, both players are aware of

the options available to the proposer. Falk et al. (2003) reported that adult responders took

into account the alternatives the other person had available to them when they proposed a

particular division of points. Specifically, responders reliably rejected unfair offers when the

proposer could have been fair. An unfair offer, based on outcome alone, was judged less

harshly and rejected by fewer responders when the proposer had no better options available.

These results with adults indicate that the mini-ultimatum game can be used to

examine whether there are developmental changes in fairness considerations. Specifically, it

allows us to examine whether there is a developmental shift from offer rejections based

entirely on outcome inequity, reflecting only a basic aversion to inequity, to rejections that

incorporate intentionality fairness assessments, which reflect a sensitivity to what other

options were available to the proposer and therefore whether inequity was deliberate. The

idea that the basis of fairness-related decisions changes developmentally is compatible with

the general idea of an "outcome to intention" shift in children's moral judgments, which has

been posited in literature on moral development (e.g., Cushman, Sheketoff, Wharton, &

Carey, 2013; Helwig, Zelazo, & Wilson, 2001; Zelazo, Helwig, & Lau, 1996). We are aware

of seven studies to date that have used the mini-ultimatum game with children and

adolescents (Bueno-Guerra, Leiva, Colell, & Call, 2016; Gummerum & Chu, 2014; Güroğlu,

van den Bos, & Crone, 2009; Güroğlu et al., 2011; Pelligra, Isoni, Fadda, & Doneddu, 2015;

Sutter, 2007; Wittig, Jensen, & Tomasello, 2013).

In order to examine how the relative importance of outcomes and intentions for

economic decisions develops with age, Sutter (2007) instructed children (aged between 7 and

10), adolescents (aged between 11 and 15), and adults (aged between 19 and 25) to play the

proposer as well as the responder in the mini-ultimatum game. The participants reported their

hypothetical offers and acceptance thresholds on a decision card, and did not interact with a

real partner (the so-called "cold play"). In the context of bargaining games, anonymity is an

important consideration, as personal features such as name, social distance, and attractiveness

of the participants have been shown to influence the decision-making process (e.g., Charness

& Gneezy, 2008; Marchetti, Castelli, Harlé, & Sanfey, 2011). The number of points obtained

in the game were converted into monetary reward (10 points = 2 euros). Sutter (2007) found

that, across all ages, participants showed evidence of both outcome- and intention-based

fairness assessments. Importantly, however, children and teenagers generally rejected

unequal offers more often than adults, leading Sutter to conclude that, for these groups, fair

outcomes are comparatively more important than the perception of fair intentions. Consistent

with Sutter's findings, Pelligra et al. (2015) also found that 10-year-olds varied their

rejections in a manner that suggested that they were tracking the fairness of intentions.

However, their study did not compare different age groups, so it does not shed light on

whether the tendency to incorporate intentions into fairness-related decision making increases

developmentally. In fact, the main interest in Pelligra et al.'s (2015) study was a comparison

with a group of individuals with Autism Spectrum Disorder, who failed to adjust their

rejections in the same way, providing some support for the idea that such modulation does

indeed result from using perspective-taking to track intentions.

To make the mini-ultimatum game more easily understandable to children, Güroğlu et

al. (2009) presented the task on a computer screen with visual representations of reward

quantities that could easily be compared. In line with Sutter (2007), Güroğlu et al. observed

an age-related increase in intention-sensitivity. However, unlike Sutter (2007), these authors

found no evidence that their group of children (9-year-olds) based their rejections on

intention-based fairness: Their acceptance rates did not vary as a function of the alternative

offer available to the proposers. Twelve-year-olds did show evidence of considering

intentions, but their rejection rates were less differentiated as a function of the alternative

offer than those of 18-year-olds. A very similar pattern of findings is reported by Gummerum

and Chu (2014), who found that 8-year-olds consistently based their rejection decisions on

outcome-fairness alone, with consideration of intentions increasing across adolescence,

consistent with Güroğlu et al.'s (2009) conclusion that a complete grasp of intentions in

social interactions cannot be expected until early adulthood (see also Güroğlu et al., 2011;

Güroğlu, van den Bos, & Crone, 2014).

Two studies have focused on younger children, and have attempted to establish

whether children do show some evidence of sensitivity to what choice alternatives were

available to a proposer. Wittig, Jensen, and Tomasello (2013) developed a more naturalistic

version of the mini-ultimatum paradigm in order to test pre-schoolers; they reported that

although 5-year-olds consistently rejected inequitable offers (3 gummy bears for the proposer

but only one for the responder) when the proposer could have made a fair choice (even split

of four sweets), they failed to adopt a more context-specific notion of fairness, meaning that

their rejections did not vary depending on whether the alternative, unchosen offer was also

unfair or was a generous one. Wittig and colleagues (2013) concluded that pre-school

children seem to act according to the simple principle that fair means equal that is not

moderated by a sensitivity to fairness intentions. Finally, in Bueno-Guerra et al.'s (2016)

study, 6- and 10-year-old children played the mini-ultimatum game anonymously against a

real opponent for one trial only. Bueno-Guerra et al. (2016) did not find consistent evidence

for intention sensitivity in the rejection patterns of either age group; however this may in part

have been because of the relatively small numbers of data points they had for each trial type.

Taken together, previous studies of the mini-ultimatum game conducted with children

and adolescents consistently suggest that the tendency to base rejection decisions on

intention-fairness increases developmentally. However, they do not paint a clear picture of the developmental trajectory of intention-sensitivity and inequity aversion, and in particular it is not clear when fairness-related decision making first begins to incorporate intentions.

In the current study, we re-visited the issue of developmental change in rejection patterns in a mini-ultimatum game, making three important adjustments to the paradigm. First, we tried to make the alternative, unchosen, option salient to participants. As Radke et al. (2012) point out, assessment of fairness intentions in this task involves thinking counterfactually about what outcomes would have been obtained, had the proposer chosen differently. We highlighted the alternative offer during each trial in order to make this counterfactual more noticeable to children; it could be that children are in fact potentially adept at considering intentions but tend to focus their attention on actual rather than counterfactual outcomes. In the context of other types of moral judgments, Nobes, Panagiotaki, and Engelhardt (2017) have shown that emphasizing the salience of information about intentions can assist children in shifting from outcome- to intention-based judgments (see also Nummedal & Bass, 1976). Nobes et al. (2017) argue that increasing intention salience is important because children may find it difficult to supress a more automatic response to visually apparent outcomes (Margoni & Surian, 2016).

Second, in addition to examining rejection rates in the mini-ultimatum game, we also separately explicitly asked participants to make judgments about the fairness of a variety of hypothetical offers. Whether an individual rejects an offer has been taken to be an indication of the extent to which it is viewed as fair, but rejections only indirectly measure fairness assessment, and we wanted to examine whether children's explicit judgments of fairness showed similar patterns to their rejections. Previous research has shown a discrepancy between younger children's judgments and behaviour: they define fair as equal but nonetheless have a propensity to accept unfair offers (Castelli, Massaro, Sanfey, & Marchetti,

OUTCOMES VERSUS INTENTIONS IN FAIRNESS-RELATED DECISION MAKING

2014b; see also Castelli et al., 2014a). We also asked participants to make judgments about the emotion they would feel in response to such offers. As we pointed out above, it has been suggested that the tendency to reject unfair offers may be closely linked to emotional responses to unfairness (see Zheng, Yang, Jin, Qi, & Li, 2017, for review). We were interested in whether children's reported emotions tracked both outcome- and intention-based fairness, and whether such judgments were consistent with their actual rejection behaviour. Finally, for reasons we explain in the next section, our study also examined a further factor that has not previously been examined developmentally: the role of time in fairness considerations.

Temporal Asymmetry and Fairness Assessments

It may seem irrational, but *time* can powerfully influence our moral judgments (Caruso, 2010; Eyal, Liberman, & Trope, 2008; Peetz, Gunn, & Wilson, 2010). Arguably, a morally reprehensible act should be perceived as equally reprehensible regardless of whether it happens to have taken place in the past of will take place in the future. Contrary to this expectation, an emerging body of research demonstrates that future transgressions are judged to be more deliberate, less morally acceptable, and more worthy of punishment than equivalent transgressions in the past (Burns, Caruso, & Bartels, 2012; Caouette, Wohl, & Peetz, 2012; Caruso, 2010). For example, participants who read about unfair business policies judged these to be less fair when they were told that these policies would be implemented in the future than when they were told that they had been implemented in the past (Caruso, 2010), and participants were more willing to punish a company that was planning to launch an advertising campaign targeting children than a company that had already launched this campaign (Roh & Schuldt, 2014). In another demonstration of this asymmetry in judgment, Caruso (2010) asked his participants to imagine playing an ultimatum game and being offered $1 from a $10 allocation (a split typically considered unfair and routinely rejected).

Adults imagining receiving a low offer in the future felt it was more unfair, and were more likely to reject it than those imagining playing the game in the past.

What may be driving this effect? It has been argued that people value the future more because they experience greater affect when imagining future events compared to past events (Caruso, 2010). The argument is that these emotional differences drive other asymmetries: We judge future transgressions more harshly than past ones because of the greater emotion we experience when considering them. In line with this hypothesis, previous studies found that people experience greater affect when thinking about the future compared to the affect they experience when thinking about the past, and that this effect holds for both positively and negatively valenced events (Caruso, 2010; Caruso, Gilbert, & Wilson, 2008; D'Argembeau & Van der Linden, 2004; Rasmussen & Berntsen, 2013; Van Boven & Ashworth, 2007). Furthermore, emotional ratings about events have been shown to mediate the effect of temporal location on moral judgments (Caruso, 2010).

In general, temporal asymmetries have been interpreted as evidence of a bias toward the future that may have an evolutionary basis, with it being argued that it may be adaptive to focus and care more about the future than the past (Suhler & Callender, 2012). However, there is currently very limited evidence regarding whether such biases are present from an early age, as an evolutionary account might predict. McCormack, Burns, O'Connor, Jaroslawska, and Caruso (2019) found some evidence to suggest that there may developmental increases in the tendency to spontaneously focus on the future. However, a recent study by Burns et al. (2019) found that by 6-7 years of age children reported stronger emotions when thinking about future versus past events, and that the magnitude of this asymmetry did not change developmentally across childhood and adolescence. As mentioned above, no previous developmental studies have examined whether children's moral judgments, like those of adults, are influenced by whether the acts being judged are located in

the past versus the future. It might be hypothesized that such past-future asymmetries will emerge developmentally, in line with the general idea that children and adolescents are less future-oriented than adults (e.g., Nurmi, 1991; Steinberg et al., 2009). However, if asymmetries in moral judgments stem from asymmetries in emotion when thinking about past and future acts, then Burns et al.'s (2019) findings would suggest that the former type of asymmetry may also be present from at least 6-7 years and not change developmentally. The mini-ultimatum game is a good context in which to examine this issue, because studies with adults indicate that violations of fairness trigger emotional arousal and offer rejection (Pillutla & Murninghan, 1996, Güroğlu et al., 2010; Güroğlu et al., 2011, Sanfey et al., 2003; Tabibnia et al., 2008). Thus, in line with the idea that compared to thinking about identical events in the past, thinking about events in the future arouses more intense emotion (Caruso, 2010; Caruso et al., 2008; Van Boven & Ashworth, 2007), the decision to accept or reject an offer would be expected to vary as a function of temporal location.

The Current Study

In the study reported here, children aged 6-to-8 and 9-to-11, adolescents aged 14-to-15, and adults completed two tasks: they made a series of judgments about trials in a hypothetical mini-ultimatum game and they then played the game "for real". When making judgments about trials in the hypothetical game, half of the participants were instructed to imagine that they were playing the game last week (past condition), whilst the other half were told to imagine that they will be playing the game next week (future condition). Participants had to make judgments about how likely they would be to reject various offers, how fair they judged these offers to be, and the emotion they experienced when thinking about the offer. The experimental variables of interest were the temporal location of the event (near future versus equidistant past) and the unchosen offer available to the proposer (fair alternative, equally unfair alternative, hyper-unfair alternative). Following these judgments about a

hypothetical game in the past or the future, participants were then invited to play the game

"for real" online against an anonymous proposer (in fact, the game was not genuinely real,

because the computer generated the proposals); during the game the unchosen options

available to the proposer were also varied, and we were interested in how these unchosen

alternatives affected participants' actual rejections.

The wide age-range sampled enabled us to examine whether there are developmental

changes in fairness considerations across childhood, adolescence, and into adulthood. We

were particularly interested in whether the child groups would show evidence of considering

intentions in their fairness assessments, and whether there would be an increase in the

likelihood that participants use intention-based fairness assessments. Unlike previous studies,

we were able to examine whether this was the case not just for actual rejections when

participants played the game "for real", but for participants' explicit judgments of fairness,

and their emotional reactions when imagining hypothetical offers. Moreover, because some

participants imagined hypothetical offers in the past and some in the future, we were able to

examine whether the temporal location of events had an impact on fairness assessments.

In terms of whether effects of temporal context (past versus future) would be manifest

in children's fairness assessments, on the basis of Burns et al.'s (2019) recent findings of

temporal asymmetries in emotion judgments, it might be predicted that even the youngest

children would show such effects. However, the task in question involved children having to

imagine relatively complex hypothetical past or future events, and make decisions about what

they would do in such an imaginary context, whereas Burns et al.'s study involved thinking

about a simple real event (e.g., next Christmas) and did not have a decision-making

component. Thus, it was not clear that children would demonstrate the same effects of

temporal location as adults in this type of task.

The mixed findings from previous studies made it difficult to make confident developmental predictions about whether offer context would affect judgments and rejections. Although Güroğlu et al. (2009) and Gummerum and Chu (2014) studies report an outcome-to-intention shift between childhood and adolescence in the context of an ultimatum game, a variety of studies using quite different tasks have suggested that considerably younger children can consider intentions when making at least some types of moral judgments (e.g., Cushman, Sheketoff, Wharton, & Carey, 2013; Koenig, Tiberius, & Hamlin, 2019; Margoni & Surian, 2017). Hilton and Kuhlmeier (2018) argue that task features, in particular the clarity and salience of intention information, may have large effects on whether children focus on outcomes versus intentions. We expected that by emphasizing the availability of an alternative offer in our procedure, sensitivity to intention might be evident in younger children than in previous studies using the mini-ultimatum game.

Method

Participants

The final sample consisted of 357 participants subdivided into four age categories: 97 younger children (42.27% female) with a mean age of 7 years 6 months ($SD = 8.69$ months, age range = 77 months – 106 months); 102[1] older children (45.54% female) with a mean age of 10 years 6 months ($SD = 6.23$ months) ranging between 113 months and 144 months; 90 adolescents (92.22% female) aged between 14 – 15 years ($M_{age} = 14.56$ years, $SD = 0.50$ years); and 69 adults ($M_{age} = 21.71$, $SD = 5.73$ years, age range = 17 – 42 years) 84.06% of whom were female. For ease of comparison, all mean group ages are presented in years in Table 1. Necessary per-group sample size was estimated using BUCSS R package (Anderson & Kelley, 2018; Version 1.0.0) and based on the F value and sample size reported by

---

[1] Missing values – age and gender were not reported for one participant, data from further two volunteers does not include date of birth

Güroğlu et al. (2009; Experiment 3). The required total sample size at the desired level of

assurance (.6) and statistical power (.8) assuming $\alpha = .05$ was 277.

Table 1 Participants' characteristics (mean age in years) by condition.

		N	M_{age}	SD_{age}
Younger children	Past	50	7.06	0.77
	Future	47	7.11	0.76
Older children	Past	50	10.04	0.53
	Future	49	10.02	0.66
Adolescents	Past	41	14.61	0.49
	Future	49	14.51	0.51
Adults	Past	36	21.97	6.45
	Future	33	21.46	4.91

Children and adolescents were recruited from multiple schools in the vicinity of the

first author's university (i.e., in a large UK city), and consent was obtained from school

authorities and parents or legal guardians. All participants aged 16 and younger were also

required to assent to taking part prior to testing. Adult volunteers were recruited from the

student population at the first author's university via an online booking system and provided

written consent before the experiment began. Children were rewarded with small items of

stationery for taking part in the study. Adults and adolescents received a voucher worth

between £4.50 and £5.50 (UK pounds) in return for their participation. The precise value of

the voucher was based on the number of coins collected in the online mini-ultimatum game.

Materials

This study used a modified version of the mini-ultimatum game developed by Falk et

al. (2003). All participants completed two tasks: 1) they made a series of judgments about

trials from a hypothetical game against an opponent who either had played the game a week

ago (past condition) or was going to play the game in a week's time (future condition), and 2)

they then played the game "for real" (i.e., here-and-now) game against a "real" opponent. In

fact, participants always played against the computer. In what follows, we will refer to the

first task as the hypothetical mini-ultimatum game and the second task as the "real" mini-

ultimatum game.

Hypothetical mini-ultimatum game. On each trial in the hypothetical mini-

ultimatum game, two divisions of coins were presented as a forced choice. If the responder

accepted the offer chosen by the proposer, the players received the specified amounts. If the

responder rejected the offer, both players received nothing. In all trials in the hypothetical

version of the game, the number of coins offered to the responder stayed the same; what

varied was the alternative, unchosen offer available to the proposer; see Figure 1.

Specifically, the offer selected by the proposer was always *unfair*: 8 points for the proposer

and 2 points for the responder (i.e., [8, 2] split). Hereafter the term "context" will refer to the

manipulation of alternative offers, i.e., the unselected alternatives to the unfair (8, 2) split.
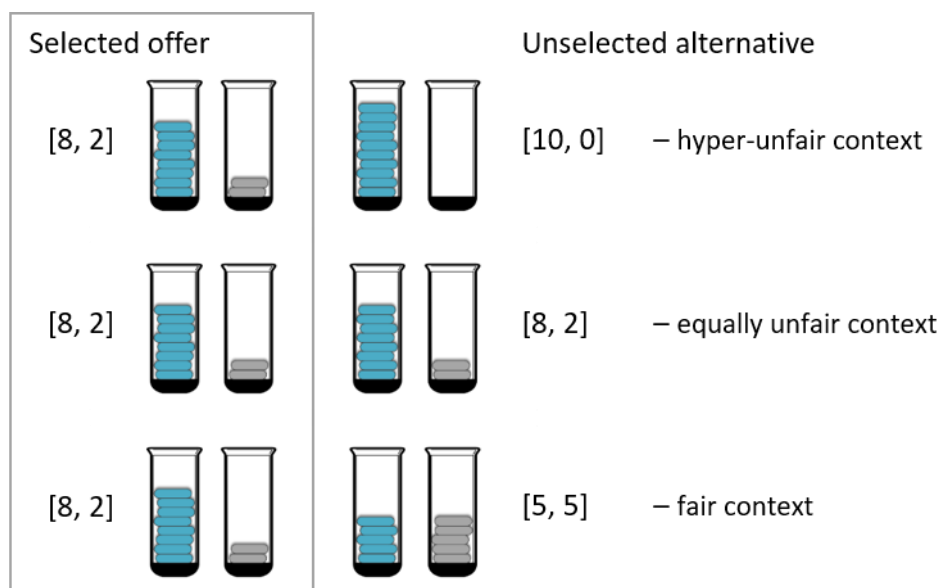
Figure 1 Stimuli used in the task in which participants made judgments about trials from a
hypothetical mini-ultimatum game. Trial context refers to the nature of the alternative but unselected
offer. Note that the proposer gets the blue coins and the responder the grey coins in each instance.

The alternative offers in the trials from the hypothetical game were: 5 coins for both

players (i.e., fair context, [5, 5] split), 8 coins for the proposer and 2 coins for the responder

(i.e., equally unfair context, [8, 2] split), or all 10 coins for the proposer (i.e., hyper-unfair

context, [10, 0] split; see Figure 1). Participants made judgments about these three trial types

only once in one of six counterbalanced orders.

"Real" mini-ultimatum game. The "real" game included the same trial types as the

hypothetical game, but also included an additional offer of 2 coins for the proposer and 8

coins for the responder (i.e., hyper-fair context, [2, 8] split). This addition resulted in four

types of trials where the fixed unfair split (8, 2) was pitted against a: fair (5, 5), equally unfair

(8, 2), hyper-unfair (10, 0), or hyper-fair (2, 8) offer. In addition, in the "real" game, there

were a further four trials in which the same options were available to the proposer but the

proposer chose the alternative rather than the (8, 2) split (see Table 2). Thus, the fixed unfair

split was chosen half of the time, and the alternative split was chosen on the other half of the

trials, resulting in 8 trials in total. Note that this manipulation resulted in two identical trials

(i.e., unfair [8, 2] paired with equally unfair [8, 2]). Using the alternative split as the chosen

offer on half the trials ensured that there was variability in the offer made to participants, and

hence made the game less repetitive. All 8 trials were presented in a randomized order. Data

analyses focused primarily on the critical trials in which the selected offer was [8, 2] (i.e., the

first four trial types in Table 2), although we provide acceptance rates below for other trial

types.

Table 2. All trials in the "real" game. Selected offers are shaded grey. Trial order was random.

	Offers available to the proposer on each of the trials		
8, 2 chosen	8, 2	or	5, 5
	8, 2	or	8, 2
	8, 2	or	10, 0
	8, 2	or	2, 8
	8, 2	or	5, 5
	8, 2	or	8, 2
	8, 2	or	10, 0
	8, 2	or	2, 8

Alternative chosen (rows 5–8)

Both tasks were programmed using E-Prime Software (Psychology Software Tools, Pittsburgh, PA). Children and some of the adults played the games on a 15inch Dell laptop with a Core i5 vPro processor and made their responses using a touchscreen. To facilitate data collection from adolescents and some adults, both tasks were re-programmed using Qualtrics (Qualtrics, Provo, UT), an online survey tool. These participants completed the task on a PC and gave responses using a mouse-click rather than on a touch-screen.

Design

Hypothetical mini-ultimatum game. The task in which participants made judgments regarding trials from a hypothetical mini-ultimatum game involved the manipulation of two variables across four age groups: first, was the unchosen offer available to the proposer (see Figure 1); second, the temporal location of the task. This resulted in a 3 x 2 x 4 mixed design with a within-subject factor of context (unchosen offers of [5, 5], [8, 2], [10, 0]) and between subject-factors of temporal location (past, future) and age group (younger children, older children, adolescents, adults).

Participants made three types of judgments per trial, meaning that there were three dependent variables: acceptance rates (*Would you have accepted this offer last week?/Would you accept this offer next week?*), judgments of fairness (*How fair was this offer?/How fair will this offer be?*), and emotion ratings (*How angry does the thought of the offer make you feel right now?*). Participants in all age groups indicated their decision to either reject or accept an offer using a 7-point scale (see Figure 2). The scale was designed to allow the participants not only to indicate their choice to either accept or reject the offer, but also to rate how certain they were of that decision. The scale ranged from *really*, through *quite sure*, to *only a little bit sure* that one would/would not accept the offer. Participants were instructed to press the middle button to indicate that they are *not sure* of their decision. Likewise, valuations of fairness (fair vs unfair) were obtained using a 7-point scale (see Figure 2)

OUTCOMES VERSUS INTENTIONS IN FAIRNESS-RELATED DECISION MAKING

ranging from *really fair/quite fair/only a little bit fair*, through *not sure*, to *only a little bit unfair/quite unfair/really unfair*. Finally, subjective judgments of emotions were given using a 7-point emotional faces scale (where 1 = *not at all angry* and 7 = *very very angry;* depicted in Figure 2). Likert-type scales have been extensively used for a wide variety of different judgment types with children aged 6 and older (e.g., Lagattuta & Sayfan, 2013; Weisberg & Beck, 2012).

       "Real" mini-ultimatum game. The "real" mini-ultimatum game employed a mixed 4 x 4 design, manipulating context (unchosen offers of [5, 5], [8, 2], [10, 0], [2, 8]) as a function of age (younger children, older children, adolescents, adults). This version of the task employed a two-alternative forced choice design. The participants had to either accept or reject the offer by clicking on one of two buttons marked with a green tick mark or a red cross, respectively.
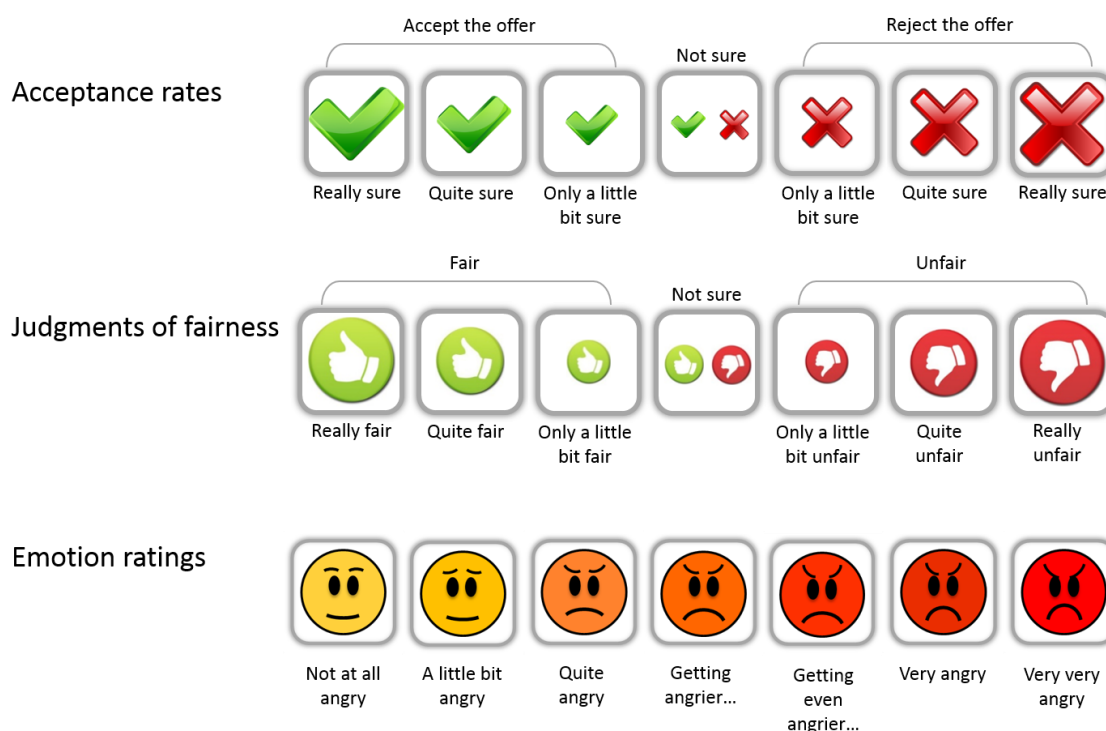

Figure 2 Response scales used in the task in which participants made judgments about trials from a hypothetical mini-ultimatum game.

Procedure

Children were assessed individually in a quiet area of the school, seated in front of a laptop PC, next to the experimenter. The experimenter guided each child through the tasks, reading aloud all verbal instructions and, if necessary, pointing to the relevant sections of the screen. Adolescents and adults using Qualtrics were tested in groups of approximately 20-30 individuals. Adults tested using E-Prime were invited to a lab at the university of the first author and were tested individually. All participants were randomly allocated to one of two temporal location conditions: past or future.

The testing session began with a training exercise designed to explain the rules of the mini-ultimatum game and to familiarize participants with the Likert-type scales used in the first task in the study. Training comprised step-by-step instructions (e.g., *There are two players in this game – The Chooser* [the proposer] *and the Decider* [the responder]) followed by a set of test questions to ensure comprehension of how the task worked (e.g., *How many coins will the Chooser get if the Decider takes this offer?*). Participants received immediate feedback and could not proceed in the game until they answered the check questions correctly, although they were allowed multiple attempts until the correct answer was chosen. The on-screen instructions were read aloud by the experimenter for child participants.

Once the central premise of the game had been explained in detail, half of the participants were instructed to imagine that they were playing the game last week (past condition), whilst the other half were told to imagine that they will be playing the game next week (future condition). The response scales were then introduced and each point on the scales was explained in turn. Participants were again required to answer test questions about each scale to ensure understanding (e.g., *Which button would you press if you think that an offer will be quite fair?*); again, the task did not proceed until participants had answered these questions correctly. The order of the tasks was fixed so that task in which participants made
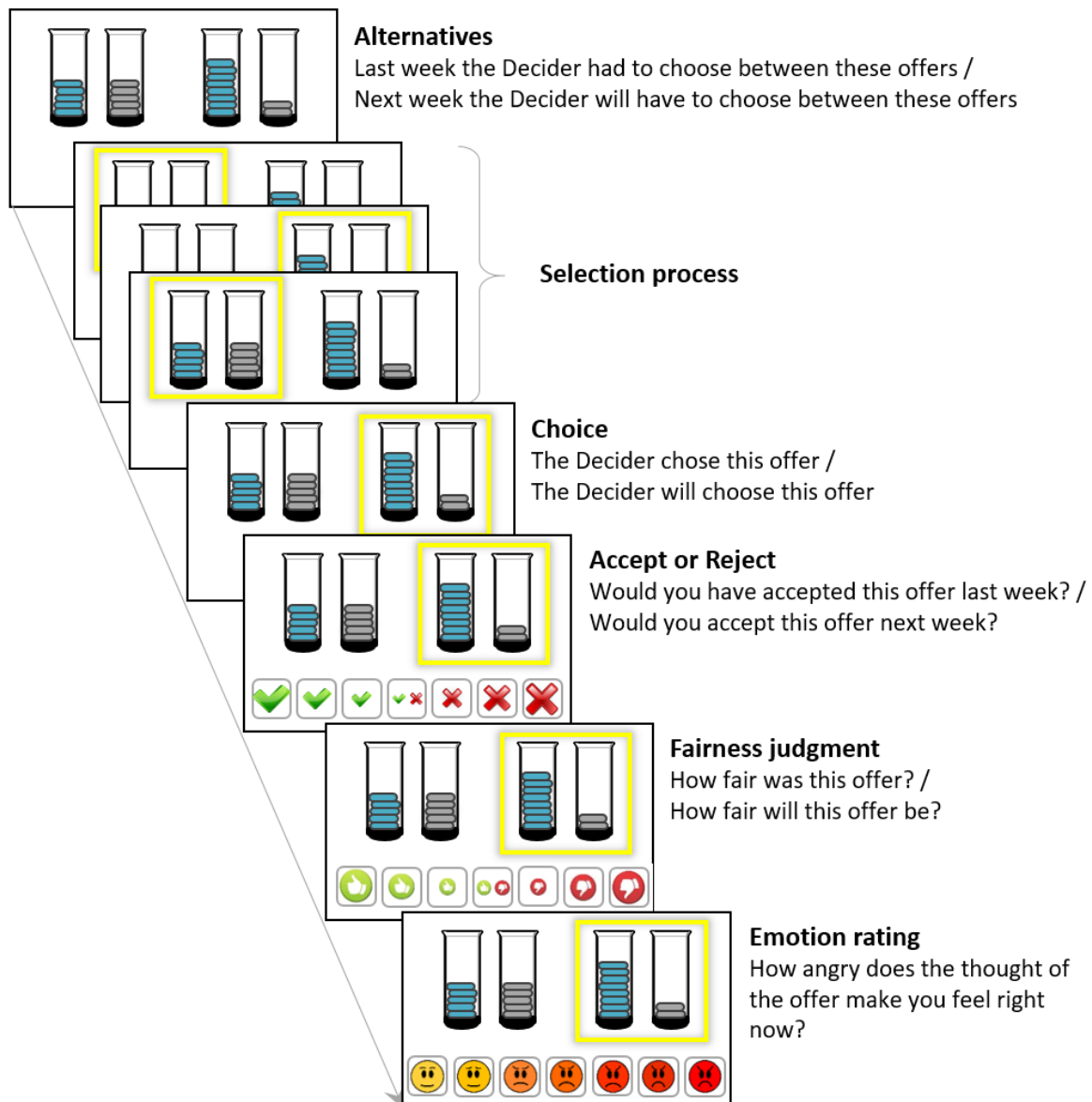
judgments about trials from a hypothetical mini-ultimatum game always preceded the task in

which participants played the game "for real". In the first task, each hypothetical trial began

with two offers (visualized as two sets of tubs containing different-coloured coins, with blue

coins for the proposer and grey coins for the responder) displayed side by side for 1000ms

(see Figure 3). Next, to emphasize that there were two available options that a proposer had

to decide between, a yellow frame flickered between the two alternatives for 650ms,

highlighting each one in turn. The proposer's choice was then revealed to the participants

(3000ms) and they were asked to answer the following three questions: 1) *Would you have*

accepted this offer last week?/*Would you accept this offer next week?*, 2) *How fair was this*

offer?/*How fair will this offer be?*, and 3) *How angry does the thought of the offer make you*

feel right now? The questions were always presented in the same order and the participants

had unlimited time to provide a response using the 7-point scales described above. Pictures of

the actual outcome and the unchosen alternative remained on the screen while participants

made their judgments.

Hypothetical game

Figure 3 The general procedure in the task in which participants made judgments about trials from a hypothetical mini-ultimatum game. To emphasize that there were two available options that a proposer had to decide between, a yellow frame flickered between the two alternatives for 650ms, highlighting each one in turn (depicted as 'Selection process' above).

Upon completion of the task involving judgments regarding a set of hypothetical

trials, the participants were invited to play the game "for real" against a random opponent.

They were then reminded of the central premise of the mini-ultimatum game (i.e., that if they

take the offer both players get coins, and if they reject the offer both players get nothing).

Children were told that if they finished with enough coins they would receive a prize; adolescents and adults were told that they would receive a reward (a voucher) between £4.50 and £5.50, depending on the number of coins they amassed across all trials in the game. Participants were then asked to pick their avatar and led to believe that they were waiting to connect with a real opponent playing in real time (see Figure 4). This was done to assure participants' motivation and to strengthen the concept of the online game.

Each trial in the "real" game began with two offers displayed side by side. Next, a yellow frame flickered between the two alternatives, to emphasize that the proposer was making up his/her mind. Finally, the proposer's choice was revealed to the participants and they were asked to either accept or reject the offer. Screen presentation durations were identical to those in the hypothetical game. All participants completed 8 trials. At the end of the experiment all children received small prizes, regardless of their performance in the online task. Voucher payoffs for adolescents and adults were contingent on task performance and varied between £4.50 and £5.50.

E-Prime and Qualtrics versions of the tasks were identical, with two exceptions. First, in the Qualtrics task the yellow frame did not oscillate back and forth between the two options. Second, the delivery of the *Choice* screen (see Figure 3 and 4) was triggered by mouse click rather than timed. For all participants testing took place in a single session lasting approximately 20 minutes. The study was approved and conducted in accordance with the guidelines of the Research Ethics Committee at the first author's university (Application number: PREC 79-2015-16). The aims of the study and methods used were communicated to parents/guardians of children ahead of data collection (in the Information Sheet). Adults and adolescents were debriefed at the end of the testing session.
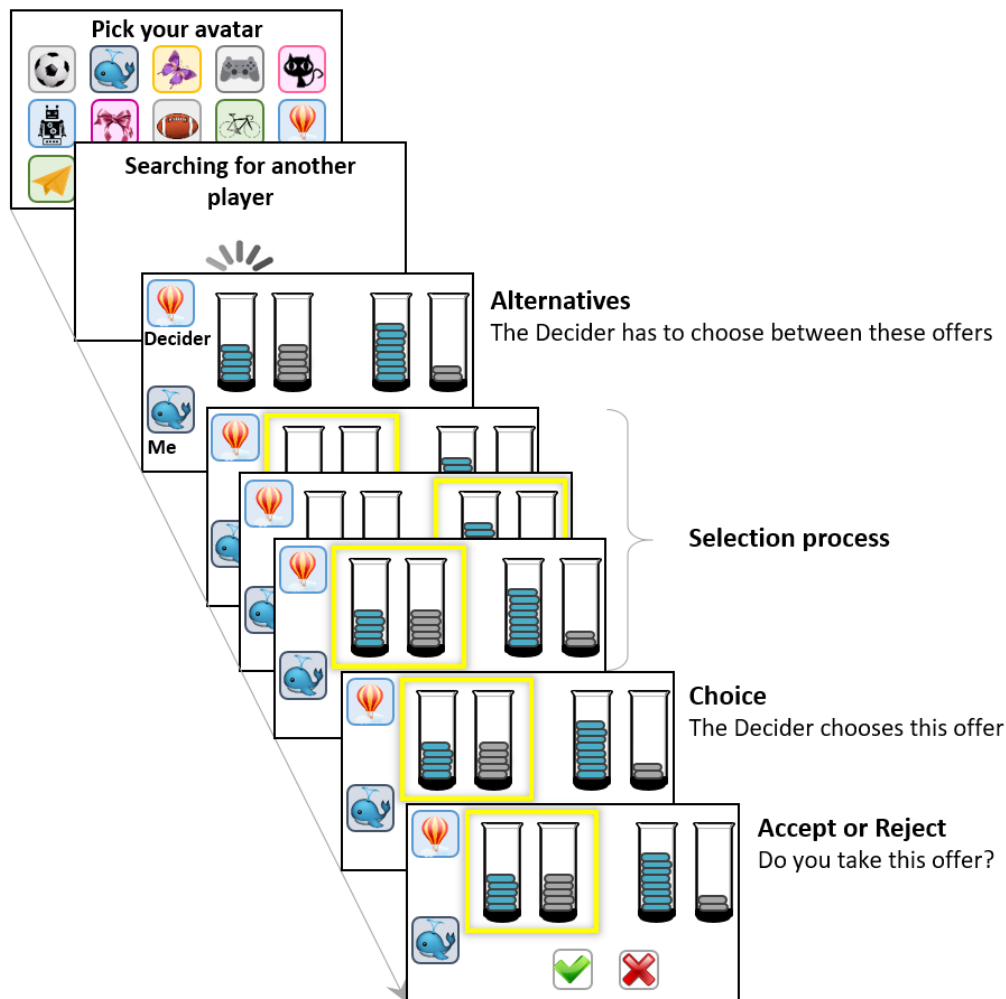
Figure 4 The general trial procedure in the "real" mini-ultimatum game. To emphasize that there were two available options that a proposer had to decide between, a yellow frame flickered between the two alternatives for 650ms, highlighting each one in turn (depicted as 'Selection process' above).

Results

Participants' Judgments About Trials From a Hypothetical Mini-Ultimatum Game

Before conducting the main analyses, we checked whether responses varied

depending on whether adult participants completed the Qualtrics or E-Prime version of the

task. A series of pairwise comparisons using acceptance rates, judgments of fairness, and

emotion ratings confirmed that the pattern of responses for adults did not vary depending on

whether they completed the E-Prime version of the task or the Qualtrics version (t(66) varied

between -1.53 and 1.78, all $ps > .08$, Bonferroni-corrected).

Figure 5 displays the participants' acceptance judgments regarding the default unfair

(8, 2) allocation across the three trials from the hypothetical mini-ultimatum game, as a

function of age group. Our primary interest was to determine whether the choice to accept the

offer was affected by the context (i.e., the alternative offer available to the proposer), age of

the participants, and temporal location (past, future). Note that temporal location was

manipulated between-participants. The data were analysed using an ordinal regression model

(cumulative link mixed model) with a logit linking function to model the log odds of

accepting/rejecting an offer on a given trial. While we plot data on its observed scale (i.e.,

percentage of judgments across all seven response categories) the modelling is done on log

odds, considered a more appropriate latent scale (see Dixon, 2008). Note also that the figures

included in the main text do not split the data by temporal location; to anticipate, this is

because analyses showed that this variable did not have a significant effect on performance

on any of the dependent measures. Figures depicting acceptance rates, judgments of fairness,

and emotions ratings split by context, age group, and temporal location can be found in the

Supplementary Materials.

The data were analysed using the *ordinal* package in *R* (Christensen, 2018; R Core

Team, 2015). Our analysis proceeded by fitting a full model, including the main effects of

temporal location, context, and age, and interactions between these variables. We used

dummy coding with equally unfair context (i.e., the available unselected offer was also [8,

2]), past temporal location, and the youngest age group as reference categories. All models

contained a random participant intercept, which modelled individual differences in overall

acceptance rates. Our approach was to fit a full model and then simplify it by removing

components and assessing change to model fit via the Schwartz's (1978) Bayesian

information criterion (BIC), which penalizes the fit of a model for the number of parameters

it has. Throughout, a positive BIC difference between two models favoured removal of an

effect/interaction, whereas a negative difference favoured retention. This process of

elimination continued until all effects not receiving support from the model comparison were

eliminated and the final model was determined. As a guide to interpreting differences in BIC

(ΔBIC) between models, Raftery (1995) suggests that a difference of ≥ 10 be considered very

strong. Where reported, the effect sizes give the magnitude of the effect in terms of expected

differences between individuals. As all models contained a random participant intercept,

estimated standard deviations were used to scale coefficients of interest.
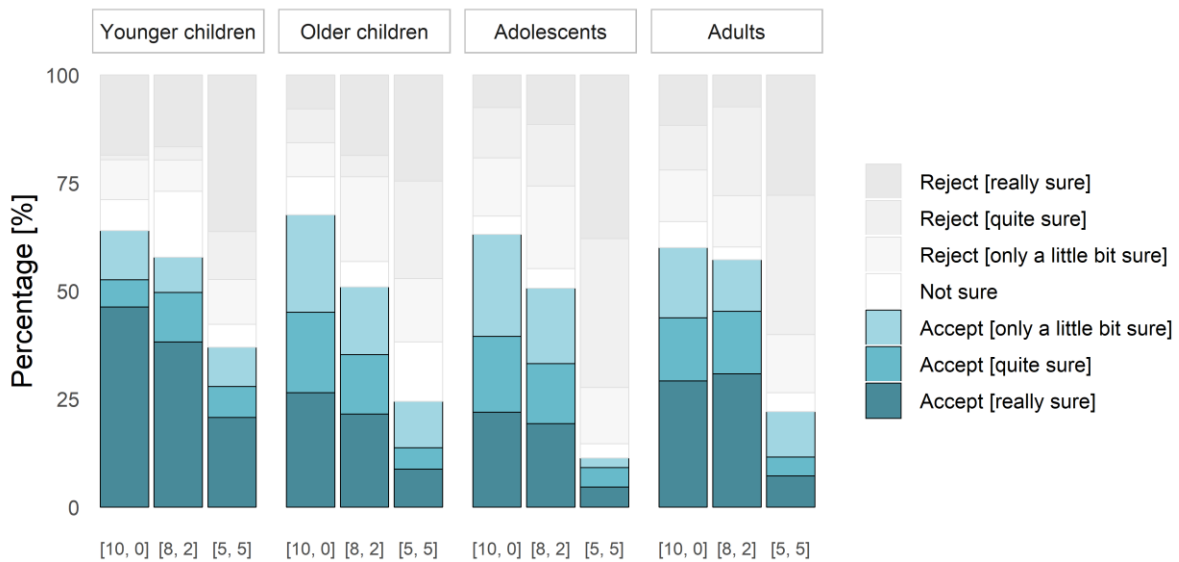
Figure 5 Participants' acceptance judgments (%) for trials from the hypothetical mini-
ultimatum game to the (8, 2) offer, split by context (the alternative offer that was available to
the proposer) and age group (younger children [aged 6-8], older children [aged 9-11],
adolescents [aged 14-15], and adults [aged $\geq$ 18]).

The final model in the analysis of acceptance rates is presented in Table 3. The full

model and description of steps taken in simplifying it can be found in the footnote to Table 3.

This final model retained the main effect of context and the main effect of temporal location.

Note that the equally unfair (8, 2) context and past temporal location were specified as

reference categories. Starting with the main effect of context, the alternative offer available to

OUTCOMES VERSUS INTENTIONS IN FAIRNESS-RELATED DECISION MAKING

the proposer had a clear effect on decision making. When context changed from equally

unfair (8, 2) to fair (5, 5), the change in the odds of moving to a higher category in the

outcome variable (i.e., from firm rejection towards firm acceptance) was 0.22. In other

words, the odds of an unfair offer presented next to an equally unfair (8, 2) alternative being

accepted are 4.55 times the odds of the same offer presented in a fair (5, 5) context. The

participant standard deviation was estimated as 1.27 so scaling this gives an effect size of -

1.20. Furthermore, the odds ratio of 1.43 (scaled effect = 0.28) indicates that, keeping all

other variables constant, as the context of the offer changes from equally unfair (8, 2) to

hyper-unfair (10, 0), the offer becomes more likely to be accepted. In line with this, the

predicted probability that a response falls into each of the seven rejection/acceptance

categories fluctuates as a function of context (reported in the supplementary material).
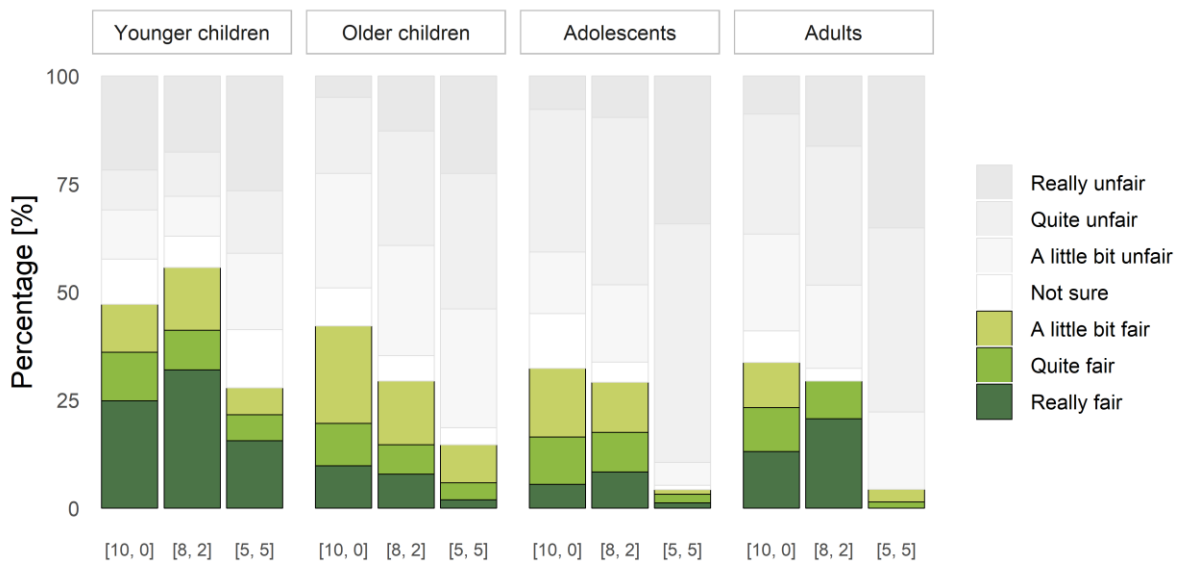
Table 3 Final model for the analysis of acceptance rates for the task in which participants made
judgments about trials from a hypothetical mini-ultimatum game.

	β (SE)	CI	Odds ratio	z
Intercept				
Reject (really sure) | Reject (quite sure)	-2.52 (0.18)	-2.87, -2.17	0.08	-13.99***
Reject (quite sure) | Reject (only a little bit sure)	-1.51 (0.16)	-1.83, -1.19	0.22	-9.19***
Reject (only a little bit sure) | Not sure	-0.73 (0.16)	-1.03, -0.42	0.48	-4.66***
Not sure | Accept (only a little bit sure)	-0.31 (0.15)	-0.61, -0.01	0.73	-2.00
Accept (only a little bit sure) | Accept (quite sure)	0.49 (0.15)	0.19, 0.79	1.64	3.18**
Accept (quite sure) | Accept (really sure)	1.23 (0.16)	0.92, 1.55	3.42	7.66***
Predictor				
Context (Fair [5, 5])	-1.53 (0.15)	-1.82, -1.24	0.22	-10.32***
Context (Hyper-unfair [10, 0])	0.36 (0.14)	0.09, 0.63	1.43	2.59*
Temporal location (Future)	-0.18 (0.18)	-0.52, 0.17	0.84	-1.00

Note The first step in reducing the full model to the one presented above was to remove the Temporal
location × Context × Age interaction term (ΔBIC = 38.8). Next, we removed the two-way interactions
(ΔBICs: Context × Temporal location = 10.1, Age × Temporal location = 20.2 and Age × Context =
32.1) and the effect of age (ΔBIC = 11.1). The effect of context was retained as its BIC comparison was
negative (ΔBIC = -170.4). Finally, the main effect of temporal location was retained because the BIC
difference of 6.2 was considered too weak to warrant omission. Confidence intervals (2.5% and 97.5%
percentile points) for parameter coefficients are provided. Significance codes: *** $p < .001$, ** $p < 0.1$,
* $p < .05$.

Figure 6 presents the participants' judgments of fairness for the default (8, 2)

allocation across three trials from the hypothetical mini-ultimatum game, split by temporal

condition for all age groups. We ran a cumulative link mixed effects model with a logit

linking function to estimate the effect of age, temporal location, and context on subjective

judgments of fairness. The final model in the analysis of fairness ratings is presented in Table

4 (see footnote for the description of steps taken in simplifying the full model).

Figure 6 Judgments of fairness for the default allocation (8,2) in the task in which
participants made judgments about trials from a hypothetical mini-ultimatum game, plotted
by context across four age groups (younger children [aged 6-8], older children [aged 9-11],
adolescents [aged 14-15], and adults [aged ≥ 18]).

The final model summarized in Table 4 retained the main effects of context, temporal

location, and age. We specified the equally unfair (8, 2) context, decisions made by younger

children, and past temporal location as reference categories. Beginning with the main effect

of context, the alternative offer available to the proposer had a discernible impact on fairness

ratings. When context changed from equally unfair (8, 2) to fair (5, 5), the change in the odds

of considering the offer to be more fair was 0.26 (participant standard deviation = 1.37,

scaled effect = -0.99). Thus, the odds of an unfair offer presented next to an equally unfair (8,

2) alternative being considered fair are 3.85 times the odds of the same offer presented in a

fair (5, 5) context. The change from equally unfair (8, 2) to hyper-unfair (10, 0) alternative

had an opposite effect on judgments of fairness (odds ratio = 1.34). With regard to the main

effect of age, in the context of an equally unfair (8, 2) alternative, younger children were

significantly more likely to judge the default (8, 2) allocation as fair when compared with all

other age groups. The predicted probabilities show that younger children are consistently

more likely to turn toward the "fair" end of the scale (see supplementary materials).

Table 4 Final model for the analysis of fairness ratings for trials from the hypothetical mini-ultimatum game.

	β (SE)	CI	Odds ratio	z
Intercept				
Really unfair | Quite unfair	-3.29 (0.25)	-3.78, -2.80	0.04	-13.14***
Quite unfair | Only a little bit unfair	-1.40 (0.23)	-1.85, -0.95	0.25	-6.11***
Only a little bit unfair | Not sure	-0.30 (0.22)	-0.73, 0.15	0.75	-1.30
Not sure | Only a little bit fair	0.17 (0.22)	-0.27, 0.61	1.18	0.75
Only a little bit fair | Quite fair	0.98 (0.23)	0.54, 1.43	2.68	4.34***
Quite fair | Really fair	1.76 (0.23)	1.30, 2.22	5.84	7.52***
Predictor				
Age (Older children)	-0.65 (0.25)	-1.14, -0.16	0.52	-2.59**
Age (Adolescents)	-1.20 (0.26)	-1.71, -0.69	0.30	-4.60***
Age (Adults)	-1.12 (0.28)	-1.67, -0.56	0.33	-3.97***
Context (Fair [5, 5])	-1.36 (0.15)	-1.65, -1.07	0.26	-9.29***
Context (Hyper-unfair [10, 0])	0.29 (0.14)	0.02, 0.56	1.34	2.10*
Temporal location (Future)	-0.08 (0.18)	-0.44, 0.28	0.92	-0.43

Note The initial step in reducing the full model to the one presented above was to remove the Temporal location × Context × Age interaction term (ΔBIC = 38.7). Next, we removed the two-way interactions (ΔBICs: Context × Temporal location = 10.8, Age × Temporal location = 19.0 and Age × Context = 21.1). The effects of temporal location (ΔBIC = 6.8), age (ΔBIC = -4.2) and context (ΔBIC = -132.8) were all retained. Confidence intervals (2.5% and 97.5% percentile points) for parameter coefficients are provided. Significance codes: *** $p < .001$, ** $p < .01$, * $p < .05$.

Post-hoc comparisons confirmed that judgments of fairness made by younger children

differed significantly from those made by all older participants ($t = 3.26$, $p_{\text{Bonferroni}} = .007$,

Cohen's $d = 0.17$, Log (BF$_{10}$) = 4.99 when compared with older children; $t = 5.14$, $p_{\text{Bonferroni}} <$

.001, $d = 0.27$, Log (BF$_{10}$) = 15.56 when compared with adolescents; and $t = 4.19$, $p_{\text{Bonferroni}} <$

OUTCOMES VERSUS INTENTIONS IN FAIRNESS-RELATED DECISION MAKING

.001, $d = 0.22$, Log (BF$_{10}$) = 8.41 when compared with adults). All other pairwise comparisons were non-significant ($p > .05$).

Anger ratings that participants gave for each trial from the hypothetical mini-ultimatum game, split by age group, are presented in Figure 7. A JZS Bayesian ANOVA with default prior scales (Rouder, Morey, Speckman, & Province, 2012) was run to determine whether the subjective anger ratings fluctuated based on the age of the participants, temporal location, and the alternative offer available to the proposer. The analysis revealed strong support for a model containing only the main effect of context relative to a null model (Log (BF$_{10}$) = 50.89) but weak support when compared to the next strongest model containing main effects of context and age and their interaction (Log (BF$_{10}$) = 1.00). Post-hoc comparisons (paired t-tests) confirmed that anger ratings were significantly higher in the presence of a fair (5, 5) alternative ($t = 9.42$, $p_{\text{Bonferroni}} < .001$, $d = 0.50$, Log (BF$_{10}$) = 36.49 and $t = 9.55$, $p_{\text{Bonferroni}} < .001$, $d = 0.51$, Log (BF$_{10}$) = 37.51, when compared with the equally unfair (8, 2) and hyper-unfair (10, 0) alternatives, respectively) but did not differ between equally unfair (8, 2) and hyper-unfair (10, 0) types of context ($t = 0.54$, $p_{\text{Bonferroni}} > .05$, $d = 0.03$, Log (BF$_{10}$) = -2.68). Anger estimates did not vary as a function of age (t-tests, all $p_{\text{Bonferroni}} > .05$, all Log (BF$_{10}$) < 3 or > -3). On average, participants reported that future unfair offers were more anger-provoking than equivalent unfair offers made in the past ($M = 3.42$ ($SD = 1.82$) and $M = 3.14$ ($SD = 1.75$), respectively), however, a post-hoc comparison revealed that the difference between past and future anger ratings was not significant, $t = -1.658$, $p_{\text{Bonferroni}} = .098$, $d = -0.09$, Log (BF$_{10}$) = 0.62.

Non-parametric correlations revealed significant relationships between anger ratings and both acceptance rates and judgments of fairness (Kendall's $\tau = -.49$, $p = < .001$, 95% CI [-0.46, -0.52] and $\tau = -.53$, $p = < .001$, 95% CI [-0.50, -0.56], respectively) indicating that more intense self-reported anger was associated with higher levels of rejection and lower

fairness ratings. The link between acceptance rates and judgments of fairness was also

significant ($\tau = .563$, $p = < .001$, 95% CI [0.59, 0.53]).
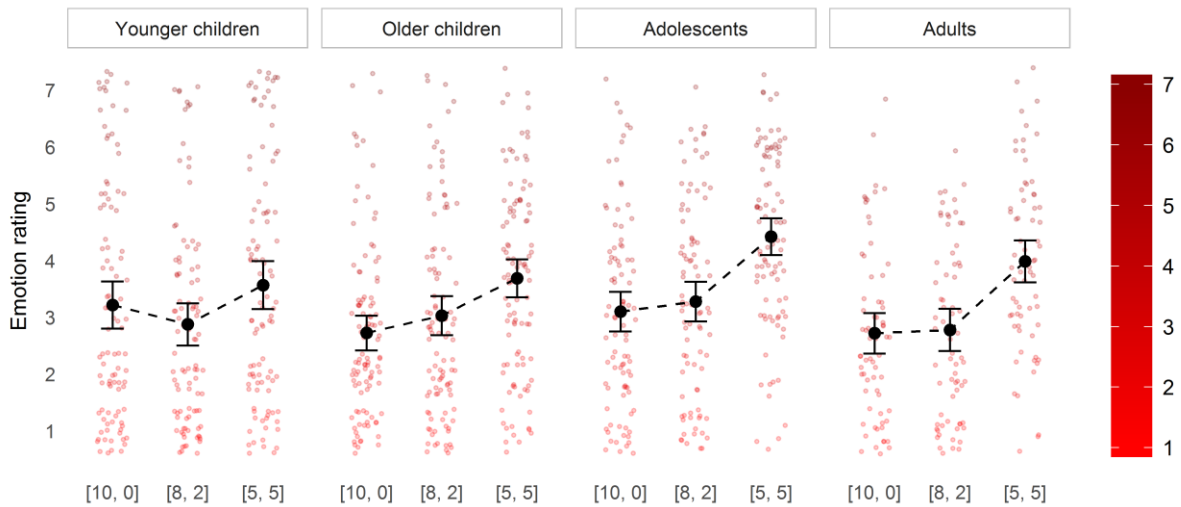
Figure 7 Anger rating for trials from the hypothetical game, split by context and age group (younger children [aged 6-8], older children [aged 9-11], adolescents [aged 14-15], and adults [aged ≥ 18]). Points are individual scores (jittered within groups to reduce overlap) with means and 95% confidence intervals overlaid.

Participants' Acceptance Rates in the "Real" Mini-Ultimatum Game

Figure 8 displays acceptance rates across all types of trials in the "real" mini-ultimatum game, plotted separately for each outcome (solid lines show trials in which the proposer selected the default [8, 2] allocation; dotted lines show trials in which the proposer selected the alternative; the "real" game included trials in which the proposer did not always select the [8, 2] offer, see Table 2), across four age groups.

Of interest here was to determine whether the choice to accept the default (8, 2) allocation (solid lines in Figure 8) was affected by the age of the participants and the alternative offer available to the proposer. Thus only this subset of the data (i.e., where the actual offer is [8, 2]) was analysed using a generalized linear mixed effects model with a logit linking function to model the log odds of accepting the offer on a given trial (using *lmer* package in R; Bates, Maechler, Bolker, & Waler, 2014; R Core Team, 2015).
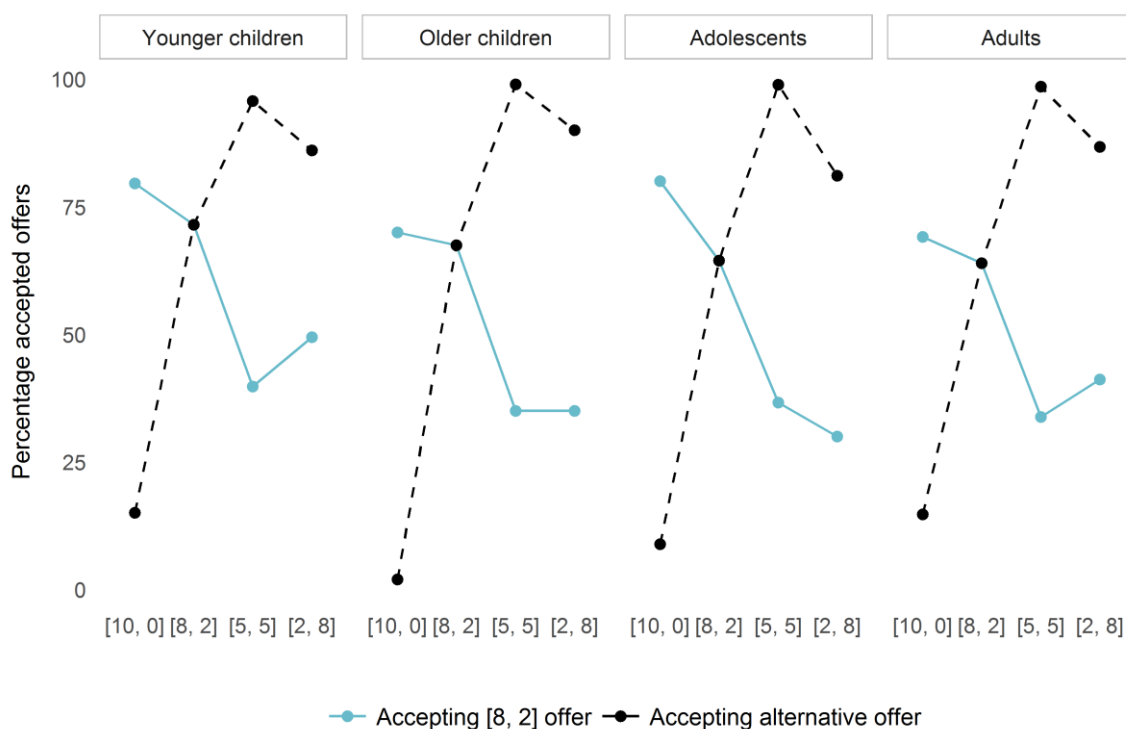
Figure 8 Acceptance rates in the "real" mini-ultimatum game as a function of type of outcome and trial type across four age groups.

The final model in the analysis of acceptance rates in the online game retained only the main effect of context (Table 5). We specified the equally unfair alternative offer as reference category. In line with the pattern of results observed in the task in which participants made judgments about trials from a hypothetical mini-ultimatum game analysis, the alternative offer available to the proposer had a clear effect on decision making. When context changed from equally unfair to fair or hyper-fair, the default (8, 2) allocations were less likely to be accepted. The participant standard deviation was estimated as 1.78 so scaling these coefficients gives effect sizes of -1.03, -1.12, and 0.33 in the hyper-fair, fair, and hyper-unfair context, respectively.

Table 5 Final model for the analysis of acceptance ratings in the online version of the mini-ultimatum game.

	β (SE)	CI	Odds ratio	z
Intercept	1.10 (0.14)	0.83, 1.40	3.02	7.64***
Context (Hyper-fair [2, 8])	-1.83 (0.18)	-2.20, -1.48	0.16	-10.09***
Context (Fair [5, 5])	-1.99 (0.18)	-2.36, -1.63	0.14	-10.75***
Context (Hyper-unfair [10, 0])	0.58 (0.18)	0.23, 0.95	1.79	3.20**

Note The first step in reducing the full model to the one presented above was to remove the Context $\times$ Age interaction term (ΔBIC = 56). Next, we removed the main effect of age (ΔBIC = 18.5). The BIC difference indicated that the effect of offer should be retained (ΔBIC = -247). Confidence intervals for 2.5% and 97.5% percentile points are provided. Significance codes: *** $p < .001$, ** $p < 0.1$, * $p < .05$.

Discussion

In this study we used a mini-ultimatum game in pursuit of two objectives. Our first goal was to examine whether there were developmental changes in the extent to which decision-making in children (aged 6-8 and 9-11) and adolescents (aged 14-15) was affected by intention-based fairness considerations. We also sought for the first time to trace the developmental origins of temporal asymmetries in judgments about fairness. To this end we tested the implications of adopting a past or a future temporal perspective on fairness-related decision-making and judgments. The key findings can be summarized as follows: Even the youngest children appear to employ a sophisticated notion of fairness that takes into account the alternatives the proposer has available to them. The acceptance data show a differential pattern of responses that varies according to the context in which an offer was made. Crucially, the likelihood of accepting an unfair split does not differ greatly by age. Across the whole sample, in the task in which participants made judgments about trials from a hypothetical game, acceptance rates were higher when participants were instructed to think about the past, compared with the future. Importantly, however, the effect of temporal location was non-significant. We discuss these findings and their implications in the sections

that follow, focussing, in turn, on context-based fairness considerations and temporal

asymmetries. We close with some prospects for future research.

The Development of Intention-Based Fairness Considerations

Our findings are in line with previous reports demonstrating that responders'

behaviour is sensitive to the alternative option available to the proposer (e.g., Falk et al.,

2003; Falk & Fischbacher, 2006; Güroğlu et al., 2009; Güroğlu et al., 2010; Radke et al.;

2012, Sutter, 2007). Specifically, acceptance rates were lower in the presence of a better

alternative (either an even split or a generous split), and higher when the default unfair split

was paired with a hyper-unfair or an equally unfair outcome. Thus, when proposers chose *not*

to offer an equal split by favouring an unfair distribution, participants seemed to recognize

the intentional social norm violation. This was true in both in the task in which participants

made judgments about trials from a hypothetical game and when participants played the

game "for real", and true for all age groups.

The most striking feature of our findings is the lack of developmental effects on

participants' rejection rates and patterns. There were no developmental changes in the

likelihood that participants would reject an unfair offer, and the pattern of rejections in

participants of all ages indicated a sensitivity to intention-based fairness. Thus, across both

mini-ultimatum game scenarios (i.e., hypothetical and "real"), the age of the participants had

no discernible effect on decision making. These findings are inconsistent with the idea that,

after the age of 7, there is a developmental shift from using outcome-based to intention-based

fairness-related decision making (Gummerum & Chu, 2014; Radke et al., 2012). Our findings

are also inconsistent with those of previous developmental studies that have used the mini-

ultimatum game with similar age ranges. At odds with Sutter (2007), we found that children

were somewhat more likely, rather than less likely, than adults to accept unequal

distributions. Higher rejection rates of unfair offers by children and adolescents reported by

Sutter (2007) were interpreted as higher inequity aversion in younger participants who perceived inequitable outcomes as more important than intentions. Furthermore, our findings do not support the claim put forward by Güroğlu et al. (2009) who argued that the tendency to incorporate a consideration of intention-based fairness into their decision-making emerges and then increases during adolescence (see also Gummerum & Chu, 2014).

Differences in methodology may account for the lack of developmental changes in our findings. Unlike Gummerum and Chu (2014) and Sutter (2007), we used displays that were child-friendly insofar as they represented the options as visual displays of quantities rather than as numbers. However, Güroğlu et al. (2009) also used quantity displays but obtained very different findings; we suspect that an important reason that our findings differ from theirs is that we had a stage in each trial in which both options available to the proposer were highlighted in turn before the decision was revealed. This meant that children's attention was drawn to the unchosen alternative as well as the actual outcome, making the counterfactual relevant to intention-based fairness considerations more salient (cf. Nobes et al., 2017). A further possibility is that children in our study were more likely to use intention-based fairness considerations because we actually asked participants explicitly to think about fairness when making judgments about trials from a hypothetical game, i.e., they had to make a series of fairness judgments. Although participants did not make such explicit judgments when playing the game for "real", it could be that making these ratings during the initial task primed children to reflect more carefully about fairness, and to base their rejections on fairness considerations that were broader than outcome inequity.

Ours is the first study to examine children's explicit fairness judgments in the mini-ultimatum game. The pattern of findings regarding these judgments was very similar to rejections, insofar as judgments regarding an (8, 2) offer were modulated by whether an equitable offer was available but unchosen, and this was true for all age groups. These data

provide direct evidence that children's judgments of fairness are affected by context rather than merely determined by the equity of outcomes, as rejection rates serve only as a proxy of such judgments. There was some developmental change in these ratings, in that, although the youngest children's fairness ratings varied across trials, in general they tended to give somewhat higher ratings (i.e., judge offers to be fairer) than older participants. This was true for all trial contexts; see Figure 6. For example, although the majority of 6- to 8-year-olds judged the (8, 2) offer to be unfair when there had been an even split available, suggesting that most children could use the fairness scale, 28% of them judged it to be at least a little bit fair. One plausible interpretation of this is that a minority of children did not properly understand the nature of the judgment that they were being asked to make or found it difficult to use the scale, which was a novel one developed for our study (only 3% of adolescents and adults judged the offer to be fair on that trial). If this interpretation is correct, then it suggests that although our findings indicate that meaningful data can be obtained from children of this age using a fairness scale, it is important to be aware that the data are likely to be noisier than those obtained with older participants. We emphasize, though, that even though it appears that a minority of children had difficulty using the scale, fairness judgments were affected by offer context in an appropriate manner in this age group, and the lack of age effects in the other measures makes it difficult to interpret these children's performance as reflecting an immature notion of fairness.

Anger ratings also varied in all age groups with offer context. Existing research with adults, using a variety of ways of measuring emotional response, including self-report, physiological, and neuropsychological measures, indicates that unfair offers in an ultimatum game provoke negative emotions in responders (e.g., Hewig et al., 2011; Pillutla & Murninghan, 1996; Sanfey et al., 2003; Vögele, Sorg, Studtmann, Weber, 2010), and our data add to this body of findings. Such findings are consistent with the so-called "wounded pride

and spite" model of ultimatum task performance, according to which rejections are assumed

to stem directly from these negative emotions (Pillutla & Murninghan, 1996; Straub &

Murnighan, 1995; Zheng et al., 2017). Our data suggest that it is at least possible that

emotions play a similar role in children's rejection decisions as they do in those of adults.

However, although anger ratings were correlated with rejection judgments in the hypothetical

game, it is not possible from our findings to infer a causal role for negative emotions, not

least because participants made their rejection judgments before they made emotion

judgments, meaning that the latter could have influenced the former. We note also that, in the

context of a one-shot ultimatum game, Steinbeis (2018) found that 6-9 year-olds' levels of

rejection of an unfair offer and self-reports of anger about the offer were differentially

affected by experimental manipulations (manipulations targeting self-control). Thus, it

remains an open question the extent to which negative emotions such as anger determine

children's rejection decisions.

Alternatives to an Intention-Based Fairness Interpretation of the Findings

In line with the standard interpretation of mini-ultimatum game performance, we have

suggested that the effects of context on children's rejections of an unfair offer (and also their

fairness/emotion judgments) reflect a sensitivity to proposers' intentions (Falk et al., 2003;

Falk & Fischbacher, 2006). However, there is some debate over whether context effects need

be interpreted in terms of intention-based fairness or whether an interpretation that does not

bring in a role for perspective-taking is plausible (Radke et al., 2012; Sandhu, 2007). Indeed,

Radke et al. (2012) argue that a distinction needs to be made between context effects per se

(i.e., the effects of an alternative offer having been possible) and intentionality effects. They

demonstrate that even when adult participants are made explicitly aware that a computer is

randomly choosing between options (i.e., choices are clearly not intentional), rejection rates

of an unfair offer vary systematically dependent on context. Radke et al. also demonstrate,

though, that rejection rates of an unfair offer when a fair offer was available are significantly higher in a condition in which participants believe they are playing against another person than when they believe the computer randomly chose the unfair offer. Taken together, these findings suggest that, although the simple availability of alternative offers modulates rejections, intentions matter over and above offer context. The idea that offer context may affect decision-making without requiring reasoning about intentions is compatible with the idea that alternative offers serve as a reference point or benchmark against which the value of the actual offer is assessed (Brandts & Solà, 2001).

Radke et al. (2012) suggest that in fairness-related decision making, taking context into consideration may emerge first developmentally and it is only subsequently that intentionality of unfairness becomes a consideration. The former is argued to depend on counterfactual thought, whereas the latter on perspective-taking abilities that improve across adolescence. In our study, in the actual game participants were told that they were playing against an online competitor in real time, and thus they could have assumed that unfair offers were deliberately chosen. However, we cannot be certain that effects of the unchosen alternative that we observed in children entirely reflect reasoning about intentions; doing so would require running a version of the task in which, as in Radke et al.'s study, participants believed that offers were randomly selected, in order to examine whether context effects occur in the absence of intentional choice. We note, though, that Radke et al.'s developmental model would suggest that context effects become strengthened with the emergence of intention-based considerations, and we found no evidence of any change in the magnitude of such effects.

In any case, these considerations suggest that future work should endeavour to extend the current findings by systematically manipulating the characteristics of the game (e.g., a peer proposer vs a computer proposer, a "one-shot" game vs repeated interactions) and

assessing their potential moderating effects on rejection rates in a developmental sample. More broadly, developmental studies of fairness-related decision making provide a context in which to examine in more detail the role of certain cognitive skills, by testing the hypothesis that specific cognitive pre-requisites must be in place developmentally before effects of context or intentionality can be observed. Specifically, researchers conducting subsequent studies could consider how children's performance on the mini-ultimatum game relates to other aspects of cognition, such as counterfactual reasoning and perspective-taking. We note, though, that fully exploring the role of such abilities is likely to involve testing a sample younger than the children included in our study. We did not include children in the preschool age range because of the likely difficulties they would have had with a variety of aspects of the paradigm, including use of our rating scales and the requirement to make a series of numerical comparisons. However, our findings suggest that future studies with younger children would also benefit from employing a methodology in which an unchosen alternative is made salient to children.

Temporal Asymmetry

We found no evidence, even in adults, that fairness-based decision making was affected by whether the events about which participants had to make a decision were assumed to occur in the past or the future. This finding contrasts with that of Caruso (2010), who found that adults judged that they would be more likely to reject an unfair offer in an ultimatum game when the game was set in the future versus the past. We also failed to find evidence that emotional reactions to unfair offers varied depending on the temporal location of the offer, whereas Caruso found that there were such temporal differences in emotions and that these differences appeared to explain the temporal differences in rejection judgments. In Burns et al.'s (2019) previous developmental study of temporal asymmetries, children made (non-moral) judgments about real rather than hypothetical scenarios, suggesting that temporal

asymmetries in judgments reflecting fairness considerations might be easier to observe for real rather than hypothetical events. It may also be that stronger emotions would be evoked if the games were not computer-based and involved a real-life protagonist, which in turn may enhance the chances of observing temporal asymmetries in fairness judgments, although we note that studies using physiological measures show that negative emotions are evoked even when the protagonist is a virtual one (e.g., Civai, Corradi-Dell'Acqua, Gamer, & Rumiati, 2010; Vögele et al., 2010).

One reason we may have not found similar asymmetries is that all the scenarios used in Caruso's experiments were simple and described briefly: Other than basic information about what had or would happen, no other information except temporal location was supplied. Participants also only made judgments once, about a single event. Our task was clearly more demanding, with participants making judgments about three separate, more complex scenarios. Because participants in our study had to process more information about the complex scenarios, and had multiple opportunities to compare outcomes, it is possible that their responses were more deliberate and thoughtful than the more quick and intuitive responses that might be elicited in a simple, one-trial procedure.

Conclusion

The findings reported here suggested that there is remarkably little developmental change across a wide age range in the basis of fairness-related decision making, at least within the context of the mini-ultimatum game. Thus, these results do not support the idea of an outcome-to-intention developmental shift in fairness-related judgments that continues across childhood and into adolescence. However, there are still unanswered questions about the development of fairness-related decision making. The youngest children in our study were 6-7 years, so it remains possible that intentions do not feature in the same way in the fairness-related decision making of pre-schoolers (Wittig et al., 2013). Moreover, even within

the age range we have studied, additional research is required to examine the role of emotion in children's fairness-related decisions, and to confirm that the pattern of effects in children's judgments that we have observed reflects the intentional component of decisional context specifically.

OUTCOMES VERSUS INTENTIONS IN FAIRNESS-RELATED DECISION MAKING

Acknowledgments

References

Anderson, S. F., & Kelley, K. (2018), BUCSS: Bias and uncertainty corrected sample size [Computer software and manual]. Retrieved from https://CRAN.R-project.org/package=BUCSS (R package version 1.0.0)

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4 [Computer software and manual]. Retrieved from http://CRAN.R-project.org/package=lme4 (R package version 1.1-7)

Beck, S. R., & Riggs, K. J. (2014). Developing thoughts about what might have been. *Child Development Perspectives*, *8*(3), 175-179.

Bicchieri, C., & Chavez, A. (2010). Behaving as expected: Public information and fairness norms. *Journal of Behavioral Decision Making*, *23*(2), 161-178.

Blake, P. R., McAuliffe, K., & Warneken, F. (2014). The developmental origins of fairness: The knowledge–behavior gap. *Trends in Cognitive Sciences*, *18*(11), 559-561.

Blount, S. (1995). When social outcomes aren′t fair: The effect of causal attributions on preferences. *Organizational Behavior and Human Decision Processes*, *63*(2), 131-144.

Brandts, J., & Solà, C. (2001). Reference points and negative reciprocity in simple sequential games. *Games and Economic Behavior*, *36*(2), 138-157.

Bueno-Guerra, N., Leiva, D., Colell, M., & Call, J. (2016). Do sex and age affect strategic behavior and inequity aversion in children? *Journal of Experimental Child Psychology*, *150*, 285-300.

Burns, Z. C., Caruso, E. M., & Bartels, D. M. (2012). Predicting premeditation: Future behavior is seen as more intentional than past behavior. *Journal of Experimental Psychology: General*, *141*(2), 227-232.

Burns, P., McCormack, T., Jaroslawska, A. J., Fitzpatrick, A., McGourty, J., & Caruso, E. M. (2019). The development of asymmetries in past and future thinking. *Journal of Experimental Psychology: General, 148*(2), 272-288.

Camerer, C. F. (2003). Strategizing in the brain. *Science*, *300*(5626), 1673-1675.

Caouette, J., Wohl, M. J., & Peetz, J. (2012). The future weighs heavier than the past: Collective guilt, perceived control and the influence of time. *European Journal of Social Psychology*, *42*(3), 363-371.

Caruso, E. M. (2010). When the future feels worse than the past: A temporal inconsistency in moral judgment. *Journal of Experimental Psychology: General*, *139*(4), 610.

Caruso, E. M., Gilbert, D. T., & Wilson, T. D. (2008). A wrinkle in time: Asymmetric valuation of past and future events. *Psychological Science*, *19*(8), 796-801.

Castelli, I., Massaro, D., Bicchieri, C., Chavez, A., & Marchetti, A. (2014a). Fairness norms and theory of mind in an ultimatum game: judgments, offers, and decisions in school-aged children. *PloS ONE*, *9*(8), e105024.

Castelli, I., Massaro, D., Sanfey, A. G., & Marchetti, A. (2014b). "What is fair for you?" Judgments and decisions about fairness and Theory of Mind. *European Journal of Developmental Psychology*, *11*(1), 49-62.

Charness, G., & Gneezy, U. (2008). What's in a name? Anonymity and social distance in dictator and ultimatum games. *Journal of Economic Behavior & Organization*, *68*(1), 29-35.

Christensen, R. H. B. (2018). ordinal - Regression Models for Ordinal Data. [Computer software manual]. Retrieved from: http://www.cran.r-project.org/package=ordinal/ (R package version 2018.4-19)

Civai, C., Corradi-Dell'Acqua, C., Gamer, M., & Rumiati, R. I. (2010). Are irrational

    reactions to unfairness truly emotionally-driven? Dissociated behavioural and

    emotional responses in the Ultimatum Game task. *Cognition*, *114*(1), 89-95.

Crone, E. A. (2013). Considerations of fairness in the adolescent brain. *Child Development*

    Perspectives, *7*(2), 97-103.

Cushman, F., Sheketoff, R., Wharton, S., & Carey, S. (2013). The development of intent-

    based moral judgment. *Cognition*, *127*(1), 6-21.

D'Argembeau, A., & Van der Linden, M. (2004). Phenomenal characteristics associated with

    projecting oneself back into the past and forward into the future: Influence of valence

    and temporal distance. *Consciousness and Cognition*, *13*(4), 844-858.

Dawkins, M. B., Sloane, S., & Baillargeon, R. (2019). Do infants in the first year of life

    expect equal resource allocations? *Frontiers in Psychology*, 10:116, 1-19.

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and*

    Language, 59 (4), 447–456.

Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind

    continues to develop in late adolescence. *Developmental Science*, *13*(2), 331-338.

Eyal, T., Liberman, N., & Trope, Y. (2008). Judging near and distant virtue and vice. *Journal*

    of Experimental Social Psychology, *44*(4), 1204-1209.

Falk, A., Fehr, E., & Fischbacher, U. (2003). On the nature of fair behavior. *Economic*

    Inquiry, *41*(1), 20-26.

Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*,

    54(2), 293-315.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The*

    Quarterly Journal of Economics, *114*(3), 817-868.

Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources. *Developmental Science*, *14*(5), 1012-1020.

Gummerum, M., & Chu, M. T. (2014). Outcomes and intentions in children's, adolescents', and adults' second-and third-party punishment behavior. *Cognition*, *133*(1), 97-103.

Güroğlu, B., van den Bos, W., & Crone, E. A. (2009). Fairness considerations: increasing understanding of intentionality during adolescence. *Journal of Experimental Child Psychology*, *104*(4), 398-409.

Güroğlu, B., van den Bos, W., & Crone, E. A. (2014). Sharing and giving across adolescence: an experimental study examining the development of prosocial behavior. *Frontiers in Psychology*, 5: 291.

Güroğlu, B., van den Bos, W., Rombouts, S. A., & Crone, E. A. (2010). Unfair? It depends: neural correlates of fairness in social context. *Social Cognitive and Affective Neuroscience*, *5*(4), 414-423.

Güroğlu, B., van den Bos, W., van Dijk, E., Rombouts, S. A., & Crone, E. A. (2011). Dissociable brain networks involved in development of fairness considerations: Understanding intentionality behind unfairness. *Neuroimage*, *57*(2), 634-641.

Güth, W., & Kocher, M. G. (2014). More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature. *Journal of Economic Behavior & Organization*, *108*, 396-409.

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, *3*(4), 367-388.

Harris, P. L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, *61*(3), 233-259.

Helwig, C. C., Zelazo, P. D., & Wilson, M. (2001). Children's judgments of psychological harm in normal and noncanonical situations. *Child Development*, *72*(1), 66-81.

Hilton, B. C., & Kuhlmeier, V. A. (2018). Intention attribution and the development of moral evaluation. *Frontiers in Psychology*, *9*: 2663, 1-5.

JASP Team (2018). JASP (Version 0.8.6) [Computer software]. Retrieved from https://jasp-stats.org

Koenig, M. A., Tiberius, V., & Hamlin, J. K. (2019). Children's judgments of epistemic and moral agents: From situations to intentions. *Perspectives on Psychological Science*, *14*(3), 344-360.

Lagattuta, K. H., & Sayfan, L. (2013). Not all past events are equal: Biased attention and emerging heuristics in children's past-to-future forecasting. *Child Development*, *84*(6), 2094-2111.

LoBue, V., Nishida, T., Chiong, C., DeLoache, J. S., & Haidt, J. (2011). When getting something good is bad: Even three-year-olds react to inequality. *Social Development*, *20*(1), 154-170.

Marchetti, A., Baglio, F., Castelli, I., Griffanti, L., Nemni, R., Rossetto, F., ... & Massaro, D. (2019). Social decision making in adolescents and young adults: Evidence from the ultimatum game and cognitive biases. *Psychological Reports*, *122*(1), 135-154.

Marchetti, A., Castelli, I., Harlé, K. M., & Sanfey, A. G. (2011). Expectations and outcome: The role of proposer features in the Ultimatum Game. *Journal of Economic Psychology*, *32*(3), 446-449.

Margoni, F., & Surian, L. (2016). Explaining the U-shaped development of intent-based moral judgments. *Frontiers in Psychology*, *7*, 219.

Margoni, F., & Surian, L. (2017). Children's intention-based moral judgments of helping agents. *Cognitive Development*, *41*, 46-64.

Margoni, F., & Surian, L. (2018). Infants' evaluation of prosocial and antisocial agents: A

 meta-analysis. *Developmental Psychology*, *54*, 1445-1455.

McCormack, T., Burns, P., O'Connor, P., Jaroslawska, A., & Caruso, E. M. (2019). Do

 children and adolescents have a future-oriented bias? A developmental study of

 spontaneous and cued past and future thinking. *Psychological Research*, *83*(4), 774-

 787.

Nobes, G., Panagiotaki, G., & Engelhardt, P. E. (2017). The development of intention-based

 morality: The influence of intention salience and recency, negligence, and outcome on

 children's and adults' judgments. *Developmental Psychology*, *53*(10), 1895.

Nummedal, S. G., & Bass, S. C. (1976). Effects of the salience of intention and consequence

 on children's moral judgments. *Developmental Psychology*, *12*(5), 475-476.

Nurmi, J. E. (1991). How do adolescents see their future? A review of the development of

 future orientation and planning. *Developmental Review*, 11, 1-59.

O'Connor, E., McCormack, T., & Feeney, A. (2014). Do children who experience regret

 make better decisions? A developmental study of the behavioral consequences of

 regret. *Child Development*, *85*(5), 1995-2010.

Peetz, J., Gunn, G. R., & Wilson, A. E. (2010). Crimes of the past: Defensive temporal

 distancing in the face of past in-group wrongdoing. *Personality and Social*

 Psychology Bulletin, *36*(5), 598-611.

Pelligra, V., Isoni, A., Fadda, R., & Doneddu, G. (2015). Theory of mind, perceived

 intentions and reciprocal behaviour: Evidence from individuals with Autism Spectrum

 Disorder. *Journal of Economic Psychology*, *49*, 95-107.

Pillutla, M. M., & Murnighan, J. K. (1996). Unfairness, anger, and spite: Emotional

 rejections of ultimatum offers. *Organizational Behavior and Human Decision*

 Processes, *68*(3), 208-224.

Psychology Software Tools, Inc. [E-Prime 3.0]. (2016). Retrieved from

    https://www.pstnet.com

Qualtrics (2016). Qualtrics software [Online survey tool]. Provo, Utah, USA. Retrieved from

    https://www.qualtrics.com

R Core Team. (2015). R: A Language and Environment for Statistical Computing [Computer

    software manual]. Vienna, Austria. Retrieved from http://www.R-project.org/

Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American*

    Economic Eeview, 1281-1302.

Radke, S., Güroğlu, B., & de Bruijn, E. R. (2012). There's something about a fair split:

    Intentionality moderates context-based fairness considerations in social decision-

    making. *PLoS One*, *7*(2), e31491.

Rafetseder, E., Cristi-Vargas, R., & Perner, J. (2010). Counterfactual reasoning: Developing a

    sense of "nearest possible world". *Child Development*, *81*(1), 376-389.

Rafetseder, E., & Perner, J. (2010). Is reasoning from counterfactual antecedents evidence for

    counterfactual reasoning? *Thinking & Reasoning*, *16*(2), 131-155.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological*

    Methodology, *25*, 111–163.

Rasmussen, A. S., & Berntsen, D. (2013). The reality of the past versus the ideality of the

    future: Emotional valence and functional differences between past and future mental

    time travel. *Memory & Cognition*, *41*(2), 187-200.

Roh, S., & Schuldt, J. P. (2014). Where there's a will: Can highlighting future youth-targeted

    marketing increase support for soda taxes? *Health Psychology*, *33*(12), 1610-1613.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes

    factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356-374.

Sandbu, M. E. (2007). Fairness and the roads not taken: An experimental test of non-reciprocal set-dependence in distributive preferences. *Games and Economic Behavior*, *61*(1), 113-130.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, *300*(5626), 1755-1758.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6 (2), 461–464.

Sloane, S., Baillargeon, R., & Premack, D. (2012). Do infants have a sense of fairness? *Psychological Science*, *23*(2), 196-204.

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.

Steinbeis, N. (2018). Taxing behavioral control diminishes sharing and costly punishment in childhood. *Developmental Science*, *21*(1), e12492.

Steinberg, L., Graham, S., O'Brien, L., Woolard, J., Cauffman, E., & Banich, M. (2009). Age differences in future orientation and delay discounting. *Child Development*, *80*(1), 28-44.

Straub, P. G., & Murnighan, J. K. (1995). An experimental investigation of ultimatum games: Information, fairness, expectations, and lowest acceptable offers. *Journal of Economic Behavior & Organization*, *27*(3), 345-364.

Suhler, C., & Callender, C. (2012). Thank goodness that argument is over: Explaining the temporal value asymmetry. *Philosophers' Imprint*, *12*(15), 1-16.

Sutter, M. (2007). Outcomes versus intentions: On the nature of fair behavior and its development with age. *Journal of Economic Psychology*, *28*(1), 69-78.

Tabibnia, G., Satpute, A. B., & Lieberman, M. D. (2008). The sunny side of fairness: preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychological Science*, *19*(4), 339-347.

Weisberg, D. P., & Beck, S. R. (2012). The development of children's regret and relief. *Cognition & Emotion*, *26*(5), 820-835.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development*, *72*(3), 655-684.

Wittig, M., Jensen, K., & Tomasello, M. (2013). Five-year-olds understand fair as equal in a mini-ultimatum game. *Journal of Experimental Child Psychology*, *116*(2), 324-337.

Valle, A., Massaro, D., Castelli, I., & Marchetti, A. (2015). Theory of mind development in adolescence and early adulthood: the growing complexity of recursive thinking ability. *Europe's Journal of Psychology*, *11*(1), 112-124.

Van Boven, L., & Ashworth, L. (2007). Looking forward, looking back: Anticipation is more evocative than retrospection. *Journal of Experimental Psychology: General*, *136*(2), 289.

Van Dijk, E., & Vermunt, R. (2000). Strategy and fairness in social decision making: Sometimes it pays to be powerless. *Journal of Experimental Social Psychology*, *36*(1), 1-25.

Vögele, C., Sorg, S., Studtmann, M., & Weber, H. (2010). Cardiac autonomic regulation and anger coping in adolescents. *Biological Psychology*, *85*(3), 465-471.

Zelazo, P. D., Helwig, C. C., & Lau, A. (1996). Intention, act, and outcome in behavioral prediction and moral judgment. *Child Development*, *67*(5), 2478-2492.

Zheng, Y., Yang, Z., Jin, C., Qi, Y., & Liu, X. (2017). The influence of emotion on fairness-related decision making: A critical review of theories and evidence. *Frontiers in Psychology*, *8*, 159.