Kindernomics: The Developmental Origins of Other-Regarding Preferences in Children

Liam Pollock

Thesis submitted in fulfilment of the degree of Doctor of Philosophy

School of Biological and Chemical Sciences
Queen Mary University of London

August, 2015

Statement of Originality

I, Liam Pollock, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature:
Date:

Abstract

People systematically allow others' outcomes to affect their decisions. These tendencies, known as other-regarding preferences, are irrational according to traditional models of economics, and yet their existence is increasingly well-documented. This picture, however, is unbalanced. More attention has been devoted to examining positive other-regarding preferences, behaviours which help others, than is the case with negative other-regarding preferences, behaviours which harm them. This thesis aimed to help rectify this imbalance by using economic experiments to study the emergence and development of negative other-regarding preferences, and the motivations which lay behind them, in childhood, in a sample aged from 4-13 years of age.

Experiments 1 and 2 focused upon costly punishment in a variant of the ultimatum game. Only children aged 6-7 years and upwards were observed to consistently show negative other-regarding preferences, which generally increased with age in both experiments. Experiment 3 used the moonlighting game to compare children's positive and negative other-regarding preferences, in the form of their willingness to make reciprocal responses to pro- and anti-social behaviours. Negative reciprocity exceeded positive reciprocity in children of all ages, and the two traits were not observed to be correlated within-subjects. Experiments 4 and 5 examined whether negative other-regarding preferences would undermine cooperation in two mutualistic contexts, the battle of the sex game and the stag hunt, and also in the chicken game. In all contexts, pairs of children failed to achieve cooperative outcomes.

The implications of these findings are discussed. There was strong evidence of basic fairness concerns such as disadvantageous inequity aversion and relative comparisons affecting these results, but less evidence of higher fairness concerns or of internalised standards of normative behaviour. Negative other-regarding preferences were ubiquitous throughout pre-adolescence and outstripped more cooperative inclinations in virtually all experimental contexts. Previous work may have over-estimated children's pro-social tendencies.

Acknowledgements

I would firstly like to thank my supervisors, Dr. Keith Jensen and Dr. Magda Osman, for their help, guidance, patience and support. Not only have they generously shared their expertise with me, and always been on hand to offer advice, they have also given me great freedom and encouragement to pursue my own research interests, and have trusted me to design and run my experiments as I saw fit. I could not have asked for better people to work under.

I would also like to say a particular thank you to my parents, Martin and Kristian Pollock. They have been constantly supportive and encouraging but, more than that, they have also provided me with vital practical help in everything from building experimental materials, to transport, to proof-reading. Without them the experiments presented in this dissertation would have been nowhere near as complete, for all sorts of reasons.

Also invaluable was the help of the team of students and other interns during data collection. They voluntarily gave up their time to help in the completion of this work, and did a superb job throughout. My sincere thanks to Nadine Kante, Jade Charleson, Sandi Kranenberg, Haneen Tohma, Ibrahim Suleyman, Anita Womuhai, Laura Mämmelä, Chandni Sidhpura, April Warde, Molly Sharp, Iva Bruneć, Fareeha Khan, Rhiannon Thompson and Javeria Hussein.

A huge thank you, too, to all of my friends and colleagues at Queen Mary University of London, who have been a constant source of help and, more importantly, a pleasure to work with. Thank you to Agata Ryterska, Jieyu Lv, Brian Glass and the rest of the Dynamic Learning and Decision Laboratory. Thanks too to Dean Semmens, Giula Bena, Raj Joseph, Dave Brown, Jayden van Horick, Vera Sarkol, Ben Pitcher, Luigi Baciadonna, Max Docker, Elodie Briefer, Erika Dawon, Vivek Nityananda, Kate Hunt, Erika Dawson, Martina Zoli, and many others besides.

Finally, I am also enormously grateful to all of the staff at the Science Museum, London, both for letting me use their facilities to collect data, and for their constant assistance in practical matters. My thanks to Elise Allthorpe, Corrine Burns, Lillian Babirye, and the rest of the team.

Contents

List of Tables and Figures

Chapter 1: General Introduction

Human behaviour is governed by more than rational self-interest. When we make decisions, particularly in social contexts, we routinely take the thoughts, feelings and outcomes of others into account. This tendency makes us "other-regarding", meaning that our behaviours are often affected by considerations of others. This is hard to explain from traditional evolutionary and economic standpoints which tend to assume that the individual is self-interested. Much progress has been made in recent years in understanding the reasons behind our other-regarding tendencies and the ways in which they operate. However, this progress has been skewed in its focus. Researchers have tended to focus on the prosocial side of our other-regarding tendencies, whilst ignoring the anti-social ways in which they can also cause us to act (Herrmann & Orzen, 2008). People are both hyper-cooperative and hyper-competitive when compared to other species, and yet far more research is devoted to the former topic than to the latter. This is an oversight, since without examining both it is impossible to develop a truly rounded conception of human social behaviour. Rationally, going out of your way to ruin someone's day makes no more sense than going out of your way to help them, and yet humans routinely do both of these things (Jensen, 2010).

This dissertation aims to contribute to the rectification of this imbalance by focusing directly on negative other-regarding behaviour. Specifically, it will examine its emergence in childhood. A series of eight economic experiments will be presented, with child participants drawn from an age range spanning the full breadth of childhood, from the pre-school years to early adolescence. The aim of this work is to explore when, how, and perhaps even why our unique proclivity for negative other-regarding preferences emerges and develops in the young mind. It also attempts to pose the novel questions of whether they are driven by the same mechanism as positive other-regarding preferences, and how they affect children's behaviours in cooperative contexts. Collectively, the eight studies presented here represent one of, if not the, largest attempts to date to experimentally chart the progress of negative other-regarding preferences through childhood.

1.1: Rationality and the Standard Economic Model

For much of the twentieth century, human thought and behaviour was assumed by psychologists and economists alike to be "rational". This label means that we would supposedly dispassionately weigh up the causes and benefits of an action, and then make a choice accordingly on whether to act, depending on whether or not the action suited our own material self-interest (Gigerenzer, 2008; Hertwig & Herzog, 2009). However, a deluge of research over recent decades has shown that this is not always the case. Far from undertaking mechanical cost/benefit analyses, humans routinely allow other factors to affect the degree of "utility", essentially the degree of satisfaction we derive from our decisions (Kahneman & Tversky, 1979; Tversky & Kahneman, 1986). Crucially, this sense of utility is affected by comparing our outcomes to those of others in social domains (Loewenstein, Thompson, & Bazerman, 1989). For example, one's sense of satisfaction at receiving a 2% annual raise at work might be tempered considerably by the knowledge that one's colleague had received 10%, even if one was, rationally and objectively, clearly better off afterwards than before.

Humans routinely use comparisons like this to inform our decision making in social and economic contexts. It is one of the main reasons that we have come to be universally recognised as systematic violators of the old rational-choice based model of economics (Bolton & Ockenfels, 2000; Camerer & Fehr, 2006; Charness & Rabin, 2002; Fehr & Schmidt, 1999; Hertwig & Herzog, 2009; Levine, 1998; Loewenstein et al., 1989). Particularly relevant to the current dissertation are the ways in which our unusual concern with others' outcomes causes us to act in ways which are both "hyper-cooperative" and "hyper-competitive". A full exploration, with examples, of both of these topics will be offered in sections 1.4-1.7. For now, it is sufficient to say that they refer to behaviours in which the actor voluntarily pays a net cost to help or hinder a target, even though this act is not in the actor's material interest. These two types of act, which cannot be referred to as "selfish", since they do not correspond with the actor's own best interests, are referred to by economists as "altruistic" in the case of hyper-cooperative behaviours, and "spiteful" in the case of hyper-competitive behaviours (Fehr & Fischbacher, 2002; Levine, 1998).

In acting in ways which are both inordinately pro-social and inordinately anti-social, people violate the set of assumptions on rationality outlined above, which have since come to be referred to as the "standard economic model" (SEM), adherence to which is characterised by a hypothetical rationalist known as "*Homo economicus*" (Fehr & Fischbacher, 2002; Frank, 1987; Gintis, 2000a; Henrich et al., 2001; Herrmann & Orzen, 2008; Hertwig & Herzog, 2009). The SEM, influenced by rational choice theory, essentially stated that people's decisions are governed by their own material self-interest, and that consequently, where others' outcomes

are concerned, they should be comparatively indifferent. As discussed, few if any researchers would now argue that this is the case. However, by providing a baseline standard of "rational" behaviour, the SEM serves as a useful source of theoretically coherent null-hypotheses against which new economic and psychological theories can be tested. Actions which are targeted at another and have the other's wellbeing – whether positive or negative – as their goal, as opposed one's own, are both common occurrences and clear violations of the SEM.

1.2: Other-Regarding Concerns and Other-Regarding Preferences

The purpose of my research is to examine hyper-competitiveness, and by acting as a guideline as to when competitive behaviour deviates from rational self-interest, the SEM will act as a barometer of hyper-competitive behaviour. Behaviours which differ systematically from the SEM can collectively be termed "other-regarding preferences" (ORPs), and this is another concept to which this dissertation will make frequent reference. ORPs are behaviours motivated by a social environment, which deviate from the SEM in that they cause us to expend resources to affect another individual's outcomes in a way which cannot be explained for reasons of rational self-interest. Such behaviours are motivated by "other-regarding concerns" (ORCs), emotional responses to others' outcomes which affect our own sense of utility. So, for example, the sight of another person weeping with despair may cause us to feel our own sense of sadness, regardless of whether the source of the other's grief has any practical implications for our own wellbeing. Although generally discussed in the psychological and economic literature as a positive force, ORCs and ORPs can in fact be either sympathetic or hostile towards the target actor, depending on whether or not they align with the actor's own cognitions and preferences. Positive ORCs are those which occur when the target's feelings are aligned with the actor's own, whereas negative ORCs occur when these feelings are misaligned (see Table 1).

Table 1: other-regarding concerns matrix

(adapted from Jensen, 2010; Ortony, Clore, & Collins, 1990)

	target positive feelings	target negative feelings
actor positive concern	shared happiness (aligned: +ve, +ve)	schadenfreude (misaligned: +ve, -ve)
actor negative concern	envy/resentment (misaligned: -ve, +ve)	pity (aligned: -ve, -ve)

Negative ORPs and ORCs are under-studied in the economic and psychological literature, comprising only a fraction of the research which has been expended into understanding their positive counterparts (Abbink & Herrmann, 2011; Herrmann & Orzen, 2008). This is a problem, since a clear conception of prosocial behaviour is impossible without a thorough investigation of both its enforcement mechanisms and the realistic limits of its scope. One key reason for this state of affairs may be that evolutionary theory tends to emphasise the difficulty of the emergence of cooperation, a phenomenon which does indeed require a complex series of potential assumptions to be met (Nowak, 2006), and which is still the source of much theoretical uncertainty (Kurzban, Burton-Chellew, & West, 2015). This struggle dates all the way back to Darwin himself, who recognised that explaining cooperation in a world governed by natural selection represented something of a quandary. Competition, by contrast, is simply assumed to be the natural order of life, and thus not worthy of such close examination. It is for this reason that it is important to stress here that the focus of this work is not mere competitive behaviour. That is indeed trivial to explain. Instead, hyper-competitive behaviour is that which involves behaviour which harms both target *and* actor, a criterion which makes it "spiteful" behaviour at a functional level (Jensen, 2010; Levine, 1998). Because spiteful acts incur a net cost for the actor, then, unlike regular competitive acts, they are no easier to explain in terms of the SEM than altruistic acts, yet this fact is not always given the acknowledgement which it deserves.

The focus of this dissertation will be to contribute to the rectification of this imbalance by focusing specifically on negative ORPs. Core questions will include when they emerge in ontogeny, how they develop in childhood, how they compare to positive ORPs, whether positive and negative ORPs are related mechanisms within-subjects, and how negative ORPs can affect and undermine children's willingness and ability to cooperate. In doing so, this project aims to shed light on some of the less edifying aspects of human social behaviour, but aspects which are essential to its understanding, nonetheless.

1.3: Developmental psychology and behavioural economics

In all of the experimental chapters in this dissertation, the subjects will be children aged from early childhood through to early adolescence. Using children as subjects has strong theoretical benefits in that it allows exploration of the ontogenetic origins and development of target behaviours. In the case of social, other-regarding behaviours this means that developmental studies can shed light on the age at which such tendencies start to become dominant, and how

this change corresponds to the growing mind's wider cognitive and moral development (Rosati, Wobber, Hughes, & Santos, 2014). At the youngest end of the age spectrum, it is particularly interesting to observe the types of behaviours which occur "intuitively" to children, as children of a pre-school and early school age are less socialised than older children, and certainly adults (Kurzban et al., 2015). They are therefore subject to virtually all of the same social instincts as adult subjects, yet they are less socialised into patterns of "correct" behaviour, and thus more prone to acting egotistically upon their desires than are older children and adults (Kogut, 2012; Smith, Blake, & Harris, 2013; Steinbeis, Bernhardt, & Singer, 2012).

Children begin to acquire a sensitivity to how others perceive them and a corresponding desire to maintain a good reputation at around 5-years of age (Engelmann, Herrmann, & Tomasello, 2012; Leimgruber, Shaw, Santos, & Olson, 2012). However, this awareness takes time to develop, and does not mature until much later in childhood (Rutland, Cameron, Milne, & McGeorge, 2005; A. Shaw et al., 2014), meaning that it is generally safe to assume that child participants are less concerned than adults with maintaining their reputation as normative actors. As a result, any ORPs that children do exhibit are more likely to be a direct enactment of their concerns, as opposed to simply the paying of lip-service to well-known fairness norms in order to maintain a veneer of morality. The latter is often suggested as a confounding factor in economic experiments with adults (Lamba & Mace, 2010).

A further interesting feature of children's cognition is that their reasoning is in some respects less intuitive and more methodical than that of adults. Heuristics are essentially cognitive shortcuts which allow us to intuitively make decisions by gist-knowledge or gut-feeling, without having to engage in a lengthy cost-benefit analysis every time we wish to get something done. Generally, they are very efficient, and consequently adults routinely use them in place of methodical, rational thought (Gigerenzer, 2008). Intriguingly, however, it would appear that the same is not always true of young children, since the process of heuristic thinking has to be learned, like any other (Jacobs & Klaczynski, 2002; Kanngiesser, Rossano, & Tomasello, 2015; Lagattuta, 2014; Reyna, 2012; Reyna & Ellis, 1994). Although children do use heuristics from as early as their pre-school years (Johnston, Johnson, Koven, & Keil, 2015), they do not employ them with the same strength and frequency as older subjects. Indeed the frequency with which they are deployed seems to increase steadily during childhood from the pre-school years to adolescence (Lagattuta, 2014; Reyna & Ellis, 1994).

This means that children are more likely than adults to methodically consider their circumstances, in exactly the way that the SEM predicts that all people should do, but the majority of people actually do not do. Adults, whether for reasons of socialisation or otherwise, seem to behave more pro-socially when forced to rely upon their intuitions, and less pro-socially when encouraged to reflect in more depth upon a problem. A recent study by Rand, Green and Nowak (2012) found that subjects in economic experiments became more generous when forced to make their decisions quickly and unthinkingly, and stingier when encouraged to take their time and reflect, thus suggesting that generosity is the intuitive behaviour of most participants. For children, simply relying on an ingrained pro-social intuition seems less likely. Their greater tendency to reflect fully upon dilemmas presented to them is, like their lack of socialisation, another reason why their decisions are likely to reflect real preferences as opposed to learned behaviours.

For all of the above reasons, this dissertation will attempt to draw comparisons between the behaviour of adults and that of children by applying specially adapted experiments from behavioural economics and game theory to children across an age spectrum spanning from pre-school to early adolescence. Such experiments, generally referred to as "games", as they typically involve a scenario whereby two or more participants' outcomes are dependent upon the behaviour of one or all of the actors involved, routinely provide reliable evidence of ORPs. Indeed, the complete collapse of the SEM as a viable predictor of real world human behaviour is primarily attributable to reliable findings of other-regarding tendencies across thousands of experiments using hundreds of different economic games (Camerer, 2003).

As Gummerum, Hanoch & Keller (2008) note, applying such economic experiments to children is particularly advantageous, since in most cases adults' preferences in these contexts are well-known. This means that the measuring of preferences on a given task at different ages can allow for valid quantitative comparisons as to when and how these preferences develop. In the context of the present dissertation, applying economic games to a large sample with an age range spanning the whole of pre-adolescent childhood will allow for meaningful conclusions to be drawn as to how the process of "irrational" economic behaviour (or "other-regarding" behaviour in the terminology of the current dissertation) develops, and how it affects children's behaviours and cooperative abilities in social situations.

Charting the emergence and presence of other-regarding concerns and preferences also allows for inferences to be drawn as to some of the proximate mechanisms which underpin them. In this dissertation I will focus on two of these in particular, namely fairness preferences and

norms (see sections 1.8 and 1.9 respectively). First however, it is instructive to review the existing literature on other-regarding behaviour, how it deviates from the SEM, and how it is expressed in both adulthood and childhood.

1.4: Evidence of Positive Other-Regarding Preferences in Adults

Before examining some of the proximate causes of our deviations from the SEM, it is worth quickly reviewing some of the behavioural evidence for their existence, in order to provide examples of the types of behaviours for which the experimental chapters will be searching. Adults exhibit a wide range of these behaviours in many different scenarios. Here I provide an overview of the ORPs demonstrated by typical adults, and in so doing consider some concrete experimental examples of both hyper-competitive and hyper-cooperative behaviour.

Whilst it is true that in a typical economic experiment a significant proportion of the population can be expected to behave "selfishly", which in the parlance of behavioural economics means that they adhere to the SEM (Falk, Fehr, & Fischbacher, 2008; Levine, 1998), this subset tends to be in the minority. Even in experiments such as the public goods game, in which not cooperating is the "rational" thing to do, a review of the literature shows that – initially at least – only a minority of participants do this, a highly robust pattern which occurs across all cultural groups thus far surveyed (Chaudhuri, 2011; Gächter , 2014; Ledyard, 1997). Thus, whilst it is true that a minority of subjects with distinctly SEM-like preferences are likely to be *present* in any given sample, they exist in sufficiently small quantities for the SEM to be a poor model of behaviour at a macro level (Camerer & Fehr, 2006).

The most basic example of hyper-cooperative, positive other-regarding behaviour is provided by the dictator game (Forsythe, Horowitz, Savin, & Sefton, 1994), in which a participant is given a monetary endowment to split between themselves and a peer however they wish. According to the SEM, the participant should simply keep everything. In practice, however, participants generally tend to give away something in the region of 28% (Engel, 2011). It should be noted that the modal offers are 0% and 50%, the former figure once again demonstrating there is a minority within the population that does adhere to the SEM. This notwithstanding, exactly what the motivation is behind such generosity on the part of the majority of the sample who choose to make some kind of donation to the target remains a topic of debate (e.g. Burnham & Johnson, 2005; Fehr & Fischbacher, 2003). What is undeniable, however, is that such

behaviour is clearly caused on some level by the presence of a social partner, and a consideration of that partner's outcomes. In a hypothetical alternative experiment in which an actor is simply presented with free money in an asocial context and it is implied that they might want to jettison some of it, it is hard to imagine that throwing 50% of it away (or indeed any of it away) would be a common response.

An important point to add here is that positive ORPs are often subject to considerable cross-cultural variation, and therefore clearly the product of social learning to a large extent. More importantly, however, what is also striking is that ORPs appear in one form or another in all cultures thus far surveyed, a range which now includes everyone from hunter-gatherers to modern city dwellers across all six continents (Henrich et al., 2001; Henrich et al., 2005; Herrmann, Thöni, & Gächter, 2008; B House et al., 2013). This ubiquity makes ORPs a "human universal", and therefore probably something for which our species has an evolved predisposition, since the alternative suggestion that all of the world's cultures have independently invented a potentially self-defeating preoccupation with others' outcomes seems implausible. Instead, the universality of ORPs across a wide range of cultures means that they are likely an innate aspect of human cognition (Schmidt & Sommerville, 2011b), a suggestion which, if true, would certainly help to explain our hyper-cooperativeness.

1.5: Evidence of Negative Other-Regarding Behaviour in Adults

The reliable and well-documented examples reviewed in the previous section have led to the suggestion that humans are a uniquely "ultra-social ape" (Tomasello, 2014). However, humans are also hyper-competitive, and therefore ultra-antisocial as well as ultra-social. This suggestion may initially appear paradoxical, but it holds true so long as humans are willing to violate the SEM in ways that harm others as well as ways which benefit them. This section will make clear that they demonstrably are willing to do this. When discussing human behaviour at a species level it is sometimes instructive to compare it to examples from other taxa (Rosati et al., 2014), and this is a case in point. Surveys of the social behaviours of primates, humanity's closest living relatives, and in particularly chimpanzees and bonobos, our closest relatives of all, have shown that their behaviour adheres to the SEM to the letter and they consistently prioritise their own outcomes at the expense of others' (Jensen, Call & Tomasello, 2007a; Kaiser, Jensen, Call & Tomasello, 2012; Riedl, Jensen, Call & Tomasello, 2012; but see Proctor, Williamson, de Waal & Brosnan, 2013). This behaviour is "selfish" according to the terms of experimental economics, and is therefore competitive but *not* hyper-competitive, since these

species are no more likely to take the opportunity to gratuitously diminish conspecifics' allocations than they are to boost them (Burkart et al., 2014; Jensen, Call, & Tomasello, 2007b; Jensen, Hare, Call, & Tomasello, 2006).

By contrast, evidence from economic experiments in favour of humanity's hyper-competitiveness is demonstrated by many forms of systematic, spiteful behaviour. This is behaviour in which the actor pays a cost to harm another, and the harming of the target is the goal of the act, and not merely an incidental consequence or a means to a higher end. Spiteful behaviour is distinct from both "functional" punishment – so called because it has the aim of altering or reforming the target's behaviour in a way which provides long-term benefits to the actor – and opportunistic aggression during which the actor shows simple indifference to the target's outcomes (Jensen, 2010). Both of these latter two phenomena are easily explicable in rational terms, and, not coincidentally, thus extant in non-human species, including chimpanzees and bonobos (Jensen et al., 2007b; Jensen et al., 2006). Spiteful punishment, however, has thus far not been demonstrated in primates, and may be uniquely human (Jensen, Call, & Tomasello, 2007a; Jensen et al., 2006; Kaiser, Jensen, Call, & Tomasello, 2012; Proctor, Williamson, de Waal, & Brosnan, 2013; Riedl, Jensen, Call, & Tomasello, 2012).

The classic example of spiteful economic behaviour is probably the ultimatum game (UG). The UG is originally attributable to Guth, Schmittberger & Schwarze (1982), although it has been run hundreds, if not thousands of times since (Camerer, 2003). The UG set-up is essentially the same as that of the dictator game (see Section 1.4), but with the additional feature that the target (or "responder") of the offer is no longer a passive recipient. Instead, should the responder judge the offer made by the first actor (the "proposer") to be undesirable they can reject it, causing both participants to receive nothing. Should they choose to accept the offer, both leave with the amount decreed by the proposer's suggested split. The SEM, of course, predicts that the responder should accept anything greater than zero, and thus the proposer should offer the lowest amount possible. In practice, however, responders routinely reject offers they consider to be insultingly small – typically those in the region of 20% or less – and, interestingly, proposers appear to anticipate this possibility, since their offers are typically in the region of 30-40% (Gächter, 2014), a figure considerably higher than that generally seen in the DG. A stingy proposer offer which is rejected by an unimpressed responder provides a nice example of the difference between competitiveness and hyper-competitiveness. In offering the responder a small amount and trying to maximise his or her own income, the proposer is engaging in a competitive, selfish act. However, by rejecting the offer and ensuring that both

parties leave with nothing, the responder eliminates the potential income of both parties, and thus engages in a hyper-competitive, other-regarding act.

Such behaviours on the part of responders are clearly the product of a concern with the other's outcome via social comparison, since in the absence of the latter there is no plausible explanation for the responder's rejection of free money. Social comparisons are a ubiquitous feature of human social cognition (Festinger, 1954; Fiske, 2011; Mussweiler, 2003), and the sense of envy caused by having less than another can be a prime motivator of harmful behaviours towards the person in question (Fehr, Glätzle-Rützler, & Sutter, 2013; Fiske, 2011). Indeed, people's willingness to harm those who have more than them is certainly not limited to the ultimatum game, and can be particularly pronounced when they feel that the target's wealth is somehow "undeserved" (Zizzo & Oswald, 2001).

Costly rejections of low offers in UGs have so far been found across almost every culture surveyed, although there is a great deal of variation, and the frequency of their expression amongst some hunter-gatherer societies is close to negligible (Henrich et al., 2005). Nevertheless, the overall picture suggests a species which is unusually predisposed towards spite as well as kindness. Some of the potential explanations behind such behaviours will be explored in sections 1.8, 1.9 and 1.10. First, however, it is necessary to examine other-regarding behaviour in the developing child, the key experimental focus of this dissertation.

1.6: Evidence of positive other-regarding behaviour in children

Human physical and cognitive development is an unusually slow process (Bogin, 1990; Kuzawa et al., 2014; Sheskin, Chevallier, Lambert, & Baumard, 2014). It consequently takes a long time for the growing child to be fully proficient in many of the cognitive skills necessary for social comparisons, and fully-rounded other-regarding concerns and preferences. Examples of relevant skills which have been suggested as being used for keeping tabs on social interactions include, but are by no means limited to, theory of mind, episodic memory, metacognition, reward prediction, the ability to tolerate reward delay, and just about every conceivable form of mathematical ability (Sally & Hill, 2006; Wischniewski, Windmann, Juckel, & Brüne, 2009). All of these take time to emerge and mature in the developing brain (Dumontheil, Apperly, & Blakemore, 2010; A. M. Moore & Ashcraft, 2015; Reyna, 2012). It is also important to note that the prefrontal cortex, the part of the brain which processes advanced social behaviour and is

unusually developed in humans, is amongst the last parts of the brain to fully mature, an occurrence which does not happen until early adulthood (P. Shaw et al., 2008; Steinbeis et al., 2012).

It is therefore clearly unremarkable that young children's interpretation of the social sphere and accompanying ORPs differ substantially from those of adults, and evolve substantially during childhood. There nevertheless exists compelling evidence that both humans' abilities to make social judgements and comparisons, and their accompanying concern for others' outcomes begins very early in life and develop with remarkable speed. Crucially, significant strides are made during infants' pre-verbal phase of development, suggesting that they cannot be solely attributable to the social learning of cultural norms. Prominent in exploring this pre-verbal phase have been experiments measuring violations of infants' expectations by examining how experimentally manipulated scenarios affect their staring behaviour. Via such methods Hamlin, Wynn & Bloom (2007) showed that by as early as 6-10 months infants are capable of making social evaluations, in that they appear to prefer animated figures who behave helpfully to those whose behaviour is neutral or unhelpful. By 15 months, children expect resources to be distributed equally amongst strangers (Schmidt & Sommerville, 2011a; Sloane, Baillargeon, & Premack, 2012). The second year of life also sees the beginnings of an active desire to collaborate with and help others emerge in children's behaviour. Warneken & Tomasello (2006) found children aged 18-months were willing to spontaneously aid adults in a wide range of scenarios, such as passing them desired objects which were just out of reach. This behaviour shows both an understanding that others have goals, and a desire to help them attain these goals. It also means that they are capable of identifying specifically what another actors wants, and how this can achieved with their help. Furthermore, at this age, children appear to find helping an intrinsically motivating act in its own right, and will engage in it at least as eagerly for no reward as they will for a material incentive (Warneken & Tomasello, 2014).

Children, then, possess all of the building blocks for pro-social acts, as well as an understanding that their behaviour can have a positive impact upon others' outcomes, from very early in development. These abilities seem to correspond with a desire to instigate joint actions towards goal achievement. Callaghan et al (2011) found children aged 17-20 months to be capable of collaborating with adults in joint actions and, importantly, replicated this finding in three markedly different cultural contexts (India, Peru and Canada), a detail which adds weight to the suggestion that this development is a universal and naturally occurring one at around this age. Similar work by Svetlova, Nichols and Brownell (2010), examined three different types

of helping behaviour – helpful actions, emotionally supportive behaviours, and altruistic giving – in two groups of children aged, respectively, 1.5 and 2.5 years. Like Warneken and Tomasello (2006), they found action-based helping to be present by 18-months, although only the elder age group demonstrated an understanding of emotional support. Both groups struggled with acts of material altruism, however, even though in this case it consisted merely of spontaneously handing over fairly mundane items such as blankets and hairclips from a pile in front of the child to an adult actor.

This latter finding is important, since acts of genuine altruism which occur independently of the expectation of reciprocation, no matter how small, are a key aspect of hyper-cooperativeness (Gintis, 2000b). Children understand and acquire collaborative and helpful behaviour quickly, early, and cross-culturally, so much so that it seems likely that such abilities must have an innate basis (McLoone & Smead, 2014; Tomasello, 2014). These are key antecedents of the ability to make genuine sacrifices to aid others, but they do not appear to be sufficient to spontaneously induce such pro-social behaviour in most subjects until somewhat later in development (although see Schmidt & Sommerville, 2011).

This point can be illustrated by turning to the literature on economic preferences in childhood. The distinction between children's developing preferences towards costly (altruistic) and cost-free giving in social contexts has been increasingly well-documented in recent years via the combination of three simple, binary-choice economic games, each of which involves a participant choosing between two distributions for themselves and a peer. These consist of, respectively, a "prosocial game", an "envy game" and a "sharing game" (Fehr, Bernhard, & Rockenbach, 2008). The prosocial game (1-1 distribution versus 1-0) tests the participant's proclivity for cost-free helping – since their own allocation is identical in either case – which is exhibited if they actively prefer to pick the equal option over the option which awards the peer nothing. The sharing game (1-1, versus 2-0) allows the more demanding choice of costly helping, with the awarding of a resource to the peer requiring the active surrendering of half of the participant's allocation. The envy game (1-1 versus 1- 2) provides examination of children's tendencies to evaluate and act upon relative comparisons between their own lot and that of others.

In the prosocial (non-costly) sharing game, children have been shown to systematically prefer the generous option at as early as 2.5 years (Sebastián-Enesco, Hernández-Lloreda, & Colmenares, 2013), although the ability to target this decision systemically towards a prosocial partner does not appear to be present until 5-years (House, Henrich, Sarnecka, & Silk, 2013;

Sebastián-Enesco & Warneken, 2015). Note, however, that chimpanzees consistently fail to display an active preference for (or against) the prosocial option even when doing so costs them nothing (Burkart et al., 2014; Jensen et al., 2006). In other words it is chimpanzees, and not children, who display the indifference to others' outcomes which the SEM would predict. This would suggest that deviation from the SEM, and thus deviation from the pure self-interest which characterises it, is a trait which humans have evolved uniquely at a point since the two species deviated paths from a common ancestor in evolutionary history.

A landmark study by Fehr, Bernhard and Rockenbach (2008) of children aged from 3-8 years found consistent preferences for the cost-free prosocial selection in the prosocial game across this entire age range. Interestingly, however the same was *not* true of the costly prosocial choice in the sharing game (i.e. 1-1 over 2-0), with only a very small percentage (less than 10%) of 3-4 year old children choosing this option. Again, this is evidence of pre-school children struggling to make costly altruistic choices, and therefore struggling to deviate from the SEM. The proportion of children choosing the equal outcome (1-1) in the sharing game rose steadily with age, although those willing to engage in costly prosociality were still in the minority even by 7-8 years of age. Recent work using these same games in six diverse cultural settings has found that children's costly prosociality does not necessarily increase monotonically with age in the case of the sharing game, with a marked dip appearing to occur at around 4-6 years in all of the six cultures surveyed (House et al., 2013). Nevertheless, the overall broad trend emergent from such work, as well as from related studies tracking the same behaviours beyond 8-years and into adolescence (Fehr et al., 2013; Häger, 2010) is for increases in both costly and cost-free prosocial behaviour in conjunction with increasing age, certainly once children reach middle childhood at around 7-8 years.

This finding is not universal. For example, in the prisoner's dilemma, another common economic experiment (see Chapters 6 and 7), increasing age does not necessarily predict increasing prosociality during middle to late childhood, i.e. from around 6-11 years old (Lergetporer, Angerer, Glätzle-Rützler, & Sutter, 2014; Sally & Hill, 2006), although equally this effect has sometimes been observed in studies using this experiment (Angerer, Glätzle-Rützler, Lergetporer, & Sutter, 2015; Fan, 2000; Gummerum, Takezawa, & Keller, 2009). Nevertheless, if and when a relationship between age and prosociality does occur in children's economic experiments, it is almost always in the same direction, i.e. a positive monotonic relationship. Experimental examples of the inverse relationship are rare to non-existent.

We may be born with the seed of a natural capacity for understanding fairness, and perhaps even for enacting genuine altruism. Clearly, however, like most of our cognitive and social abilities these traits take a long time to mature. The majority of the literature on children's understanding of fairness, and of their other-regarding concerns generally, has tended to focus overwhelmingly on the prosocial side of such cognitions, emphasising the remarkable speed of their maturation, and contrasting it with the failure of other species of social primates to act in the same way (for reviews, see Tomasello, 2014; Tomasello, Melis, Tennie, Wyman, & Herrmann, 2012). However, to do so ignores a raft of other distinctly more anti-social ways in which human behaviour deviates dramatically from the SEM, and the limited work which has been done to date suggests that here too children show signs of a precocious ability.

1.7: Evidence of Negative Other-Regarding Behaviour in Children

Within the past decade, empirical work has begun to consistently show that spitefulness, like generosity, has deep ontogenetic roots. Developing children quickly decide to move beyond the logical constraints of the SEM and develop a taste for hyper-competitive behaviour. The previous section reviewed a series of papers in which children's developing positive other-regarding preferences were examined, as shown by the prosocial and sharing games (Fehr et al., 2008; Fehr et al., 2013; Häger, 2010; B House et al., 2013; Sebastián-Enesco et al., 2013; Sebastián-Enesco & Warneken, 2015). Also present in two of these studies, however, was a third experiment (Fehr et al., 2008; Fehr et al., 2013). In the "envy game", children had to choose between an equal allocation (1- 1) and an allocation where they received the same in absolute terms, but less in relative terms (1-2), i.e. one in which they could prosocially confer extra resources to their partner at no cost to themselves. Adults, typically, are willing to make the prosocial choice in this game (Fehr et al., 2013), thus showing a willingness to tolerate having less than another (or "disadvantageous inequity"), at least in cases where doing so does not cost them anything. Young children (3-4 years), however, prefer not to, and not until they reach 7-8 years of age does a majority of the sample begin to prefer this form of cost-free generosity (Fehr et al., 2008). Since, even at 3-4 years, the majority of Fehr et al's (2008) sample was willing to make the cost-free generous offer in the prosocial game (i.e. 1- 1 as opposed to 1-0), it must be the case that young children are averse to disadvantageous inequity, and thus already skilled at making relative comparisons. If 1-1 is considered acceptable but 1-2 is not, in spite of the fact that the child's payoff is identical in each case, then here we have clear evidence of children's sense of utility being influenced by others' outcomes, and having a knock-on effect on their observable behaviour.

This suggestion is confirmed by a series of experiments by Blake and colleagues, also focusing on samples aged from 4-8 years, but extending the work of Fehr et al (2008; 2013) by testing the willingness of subjects across this age spectrum to pay an actual, spiteful cost – in a manner not dissimilar to a UG rejection – in order to avoid inequitable outcomes. This is a more complex behaviour than the envy game, since the less pro-social choice in the following experiments involved *sacrificing resources* to avoid disadvantageous inequity, something which the sharing, pro-social and envy games do not require. In other words, in contrast to the previous examples, the work of Blake et al demonstrates actual hyper-competitive preferences, as opposed to merely competitive preferences. Even at 4-years of age, such choices took place on around a third of trials, a proportion which rose steadily throughout the age sample until it occurred in the large majority of cases in children aged 7-8 years (Blake & McAuliffe, 2011). This result has since been replicated and extended by two follow-up studies. The first showed that spiteful rejection rates were increased by the physical presence of a partner, thus confirming the social other-regarding element to the decision (McAuliffe, Blake, Kim, Wrangham, & Warneken, 2013). The second demonstrated that spite and not frustration was the key motivator behind such decisions, as shown by the fact that rejections were more likely when they affected the target's allocation as well as the actor's, as opposed to merely the latter (McAuliffe, Blake, & Warneken, 2014).

Consistent across all three of these studies is a largely monotonic relationship between increasing age and an increasing willingness to pay a cost to satisfy other-regarding fairness concerns in a way which systematically violates the SEM. Since this statement is also true of children's positive fairness concerns, it is tempting to draw direct parallels between the two and state that children's other-regard increases with age. However, in the sphere of negative ORPs, the picture appears to be complicated by another factor. All people make social comparisons and evaluate their outcomes in relative terms (Fiske, 2011; Loewenstein et al., 1989; Mussweiler, 2003), but it is beginning to emerge that this tendency seems to be exceptionally strong in early-to-mid childhood. Recent work by Bügelmayer and Spiess (2015), added a further "costly envy" game to a study involving the usual prosocial, sharing and envy games. The former involved choosing between an equal outcome (1-1) and an unequal outcome which was superior in absolute terms, yet inferior in relative (2-4). Only 5-6 year olds were tested, but over a quarter of them chose the equal distribution, which was in this case the spiteful option since it conferred fewer overall rewards upon both parties (i.e. one each, as opposed to two for the actor and four for the target).

Sheskin et al (2014) adopted a similar paradigm, again requiring subjects to make binary choices between two distributions which differed in absolute and relative terms, and applied it to a broader age sample ranging from 5-10 years. Like Bügelmayer and Spiess (2015), they found 5-6 year olds willing to pay a cost in order to eliminate their partner's relative payoff advantage. Interestingly, however, this tendency dissipates quickly with age, with all elder age-cohorts more likely to choose the option which maximises the collective payoff.

This is important as, contrary to the work of Blake et al (2011; 2013; 2015) and Fehr et al (2008; 2013), here are examples of younger children being *more* likely to violate the SEM for reasons of social comparison. This is an unusual finding, and one which shows that the likelihood of making all other-regarding decisions does not necessarily increase in tandem with age and its accompanying developments in cognition and normative socialisation. If pro- and anti-social ORPs follow different developmental pathways then this perhaps therefore also hints that they are not necessarily driven by the same mechanism, an idea which will be explored experimentally in Chapter 5.

Anti-social as it is, it nevertheless seems likely that the behaviour of children in Sheskin et al (2014) is motivated by other-regarding concerns. Recent studies by both Fehr et al (2013) and Steinbeis and Singer (2013) have charted a marked decline in the expression, both verbal and behavioural, of the negative ORPs of envy and schadenfreude as children get older. Admittedly in both cases the sample began at a slightly older age than Sheskin et al (2014), running from 7-8 years of age upwards, but the relatively linear decline in both emotions over the course of mid- to late-childhood through adolescence and towards adulthood hints that both may be present in even higher levels in children younger than 7-years. This is especially likely when placed in the context of the previously discussed findings of House et al (2013a), which showed that the behavioural expression of prosociality appears to dip to a particularly low level in children aged around 4-6 years. Overly aggressive social comparisons have been observed to decline from around 7-8-years and upwards (Blake & McAuliffe, 2011; Fehr et al., 2008; Häger, 2010; Sheskin, Bloom, & Wynn, 2014; Steinbeis & Singer, 2013).

Cumulatively, this evidence clearly demonstrates that social comparisons influence social behaviours from very early in childhood. In particular, children seem extremely resistant to situations where they receive less than their peers, and their sense of utility is clearly affected by this. However, many questions remain unanswered. In some paradigms, such as the UG, older children display steadily greater negative ORPs than younger children (e.g. Bereby-Meyer & Fiks, 2013), yet in others the opposite is observed (e.g. Sheskin et al., 2014). Clearly, then,

negative ORPs are not always caused by the same motives, nor are they something we simply learn to use more consistently the older we get. It seems their expression may also emerge in a less linear manner than is generally observed to be the case with positive ORPs (e.g. Benenson, Pascoe & Radmore, 2007; Fehr et al., 2008; Fehr et al., 2013; Häger, 2010; although see House et al., 2013a). The fact that this is the case raises the intriguing possibility that the two processes may be caused by different mechanisms, rather than just increasingly strong manifestations of a single human preoccupation with social comparison. Having outlined the types of hyper-competitive SEM-violations upon which the experimental chapters will focus, it is now necessary to speculate as to some of the motivations which may underlie them.

1.8: Reasons for Other-Regarding Concerns 1: Fairness Concerns

Repeated findings of the type thus far reviewed have led psychologists and economists to speculate as to the underlying cognitive mechanisms behind such behaviour. One key source of these reasons is a group of cognitions generally referred to as fairness preferences[1]. Trying to incorporate all such preferences into a single model runs the risk of requiring so many ad hoc addenda to make said model almost meaningless (Hertwig & Herzog, 2009). Nevertheless, many plausible candidates for widely-shared motivators of other-regarding concerns have been unearthed by such work. It is important to note that these theories are not mutually exclusive, and are in fact largely complementary to one another. Milestone papers by Loewenstein et al. (1989) and Fehr and Schmidt (1999) have showed that people are "inequity averse", meaning that they appear to have a dislike of inequitable outcomes – particularly, though not exclusively, when they are on the wrong end of the inequity. This appears to be one of our most fundamental and ubiquitous fairness concerns, and has been observed in children as young as 3-years old (LoBue, Nishida, Chiong, DeLoache, & Haidt, 2011). Crucially, throughout the lifespan, this dislike is so strong that many people will pay a cost to avoid inequitable outcomes, meaning that in certain scenarios they derive greater utility from eliminating a disparity in incomes than they do from simply maximising their own income.

Shortly afterwards, Bolton and Ockenfels' (2000) ERC (Equity, Reciprocity, Competition) model demonstrated that, as well as equity, people also have a concern with their relative as well as their absolute outcomes, and also a desire to repay others' acts of generosity or meanness in a like-for-like manner. Again, many subjects were willing to pay absolute costs in order to realise all three of these other-regarding outcomes in a manner wholly contrary to what the SEM

[1] Or sometimes, alternatively, as "social preferences", though the terms are essentially interchangeable.

would predict. An attempt by Charness and Rabin (2002) to unify and expand both of the above models further added evidence of a motivation to maximise "social welfare", meaning the total absolute amount of resources available for division. This tendency was demonstrated experimentally by subjects who were willing to give up part of their own income if it meant generating higher payoffs for a wider group of people as a whole, particularly when the target or targets were particularly badly off in relative terms.

Finally, it has been shown that we have a sensitivity to others' intentions as well as simply the outcomes of their actions; in other words, we are often willing to reward those who have *tried* to behave generously towards us, even when they ultimately fail, and vice-versa (Falk et al., 2008; Nelson, 2002). Not all of these more complex fairness concerns appear to exist as early in childhood as does inequity aversion. Children as young as 3-4 years old are certainly capable of evaluating their outcomes in relative rather than absolute terms (Fehr et al., 2008; LoBue et al., 2011). However, they do not acquire a reliable ability to target cost-free reciprocity until they are around 5.5 years of age (House et al., 2013), and their willingness to pay a cost to reciprocate remains patchy well into late childhood (e.g. Sutter & Kocher, 2007; see Chapter 5). Finally, whilst Charness and Rabin's (2002) concept of "social welfare" has been little researched in the developmental literature, it seems unlikely that many young children would be actively in favour of such behaviours given their generally limited positive ORPs, although the possibility remains open. This will be one of the issues upon which the experiment in Chapter 6 can shed light.

One thing which all of these ideas appear to share is an attempt to capture aspects of a sense of "justice" or "fairness", which most people seem to spontaneously attend to when social comparisons are salient. It is important to note, however, that the willingness to *act* on this sense of fairness, in other words to turn concerns into preferences, is subject to large variations from individual to individual and, within each individual, from one circumstance to another (Levitt & List, 2007). A sound knowledge of what is fair does not always translate into actual fair behaviour, a theme which will recur many times over the course of this dissertation.

Evidence for many of these theories can be seen in the examples reviewed in Sections 1.4-1.7. For example, if humans are indeed motivated by avoiding inequity (Fehr & Schmidt, 1999; Loewenstein et al., 1989), as well as their relative position compared to others and a desire to reciprocate others' behaviours (Bolton & Ockenfels, 2000), and the intentions behind others' actions (Falk et al., 2008; Nelson, 2002) then it is possible to make a good case for why

rejections occur in the scenario of small and miserly proposer offer in the ultimatum game, which could be said violate all four of these principles.

Not all subjects, however, attach equal weight to all of them, and some discard or ignore some of them altogether. A neat experimental variation on the UG by Falk et al (2003) demonstrates this point. In the "mini-ultimatum game" (MUG; see Chapters 3 and 4), the proposer must choose between only two possible offers and the responder knows that this is the case. A small but substantial minority – around 10-20% of responders in the original experiment (Falk, Fehr, & Fischbacher, 2003) – rejects all offers in which the responder's share is less than the proposer's, even in conditions where the experimental set-up rendered making an equal offer impossible. This demonstrates that simple inequity aversion is sufficient on its own to cause spiteful acts in a subset of subjects, a finding which has been replicated in other economic experiments (Raihani & McAuliffe, 2012; Zizzo & Oswald, 2001). Importantly, however, most responders in the MUG are far more likely to reject an unfair offer when it is the product of the *proposer's* choice, as opposed to the experimenter's, thus demonstrating that intentionality, and a desire to reciprocate it (Bolton & Ockenfels, 2000), are also important factors in determining when and whether spiteful punishment is employed.

1.9: Reasons behind other-regard 2: Norms

A second potential source of hyper-cooperative and hyper-competitive behaviours is social norms. Norms are culturally-defined rules that dictate the "correct" codes of thought and behaviour to members of a cultural group. As such they serve an important social function by aligning the individual's thoughts and behaviours with that of their wider group (Jensen, Vaish, & Schmidt, 2014) in a manner that makes them a key pillar of the social order (Gächter, 2014). They also play a central role in motivating other-regarding behaviour, both hyper-cooperative and hyper-competitive, since they provide the individual with a consistent moral yardstick against which the measure the rectitude of others' actions.

Central to their ability to do this is the manner in which people can "internalise" the norms of their culture, to the point where they become an intrinsic source of moral guidelines in forming their behaviours towards others (Gintis, 2007; Kogut, 2012; Perugini, Gallucci, Presaghi, & Ercolani, 2003; Rutland et al., 2005). Once this occurs, people are motivated to view behaving in accordance with norms as an end in and of itself, since doing so confers utility

on them by allowing them to feel positive even in situations (e.g. returning a stranger's lost wallet) in which they have actually incurred a penalty in purely material terms (Burger, Sanchez, Imberi, & Grande, 2009; Krueger, Massey, & DiDonato, 2008; Perugini et al., 2003). In emotional terms, however, this action confers upon the individual a "warm glow" (Andreoni, 1990) of self-regarding emotions such as satisfaction and pride, which can be thought of as intrinsic rewards for such behaviour. In economic terms, these positive emotions could be described as contributing to the actor's sense of utility (Gintis, 2007). Conversely, the capacity to mentally punish ourselves when we fail to meet internalised normative standards also exists, in the forms of negative self-regarding emotions such as guilt, embarrassment and shame. The success or failure of others in meeting the social norms we value can consequently cause equivalent emotions to be directed outwardly at them. They thus become socially-directed emotions, or other-regarding concerns in the terminology of this dissertation.

Emotionally-driven reactions to perceived adherences to, and violations of, the normatively "correct" behaviour are likely strong drivers of other-regarding preferences, both in economic games and in more naturalistic settings. Norms, it must be noted, vary hugely between cultures, which is no doubt one of the reasons behind the large cross-cultural variations observed in many economic studies (Chaudhuri, 2011; Gächter, Herrmann, & Thöni, 2010; Henrich et al., 2001; Henrich et al., 2005; Herrmann et al., 2008) – people with different internalised standards of behaviour will inevitably respond differently to the same stimulus in some circumstances. Such differences notwithstanding, the existence of complex social norms *of some description or another* is indisputably a universal feature of all human cultures. Norms are thus another strong candidate for a human-unique evolved trait at the heart of our extreme sociality, and the sense of other-regard which underpins it (Rakoczy & Schmidt, 2013; Tomasello, 2009).

A number of researchers have addressed this idea in their work. Chudek and Henrich's (2011) theory of "norm-psychology", for example, suggests that humans are cognitively adapted for living in a social environment governed by social norms. Thus, although the actual content of individual norms is largely arbitrary and we have no innate predilection for any one specific set of "moral" behaviours at the expense of another, we are naturally predisposed to observe, learn and internalise the norms which our culture dictates to us. There is solid experimental support for this suggestion. By just 14 months of age subjects intuitively appear to be more willing to copy the actions of those whose spoken language marks them out as ingroup members (Buttelmann, Zmyj, Daum, & Carpenter, 2013), despite the fact that the children themselves are obviously still preverbal at this age. Children quickly move beyond a mere

preference for the ingroup, and begin to show an active interest in upholding its values. By 3-years of age children will already make spontaneous objections to those committing anti-social acts such as destroying another's property (Rossano, Rakoczy, & Tomasello, 2011), or even just playing a game by the "wrong" rules (Rakoczy, Warneken, & Tomasello, 2008).

Behaviours which suggest an adherence to cultural rules and influences continue to develop throughout early childhood. Haun and Tomasello (2011) adapted Asch's famous (1951) conformity study in pre-school children and found that, just like adults, many of the children would conform and publically endorse a majority opinion even in cases where it was obviously, objectively wrong. Perhaps most tellingly of all, however, increasing experimental evidence is accruing that shows children going beyond mere conformity to norms and instead taking an active interest in enforcing them upon others via punishment (Gummerum & Chu, 2014; Kenward & Östh, 2012; McAuliffe, Jordan, & Warneken, 2015; Riedl, Jensen, Call, & Tomasello, 2015; Robbins & Rochat, 2011). This is a topic which Section 1.10 will explore in greater detail.

Norms are a universal feature of all global cultures both present and historical, and children seem predisposed to acquire them at a rapid rate. Indeed, arguably the very conception of an identifiable culture is impossible without a set of distinct norms for it to rest upon. Chudek and Henrich's (2011) argument thus seems a plausible one. If it is true, then it also follows that norms must be one of the key forces which cause us to behave in an other-regarding manner, since they shape the way we think in almost all social contexts including, but certainly not limited to, what is and is not fair. The wide array of different pro- and anti-social acts committed by people of different cultures in economic experiments is good evidence of this (Henrich et al., 2001; Henrich et al., 2005; Henrich et al., 2006; Herrmann et al., 2008).

As an interesting aside, however, it is worth noting that there are those who are prepared to go beyond this premise and suggest that we are genetically predisposed to consider certain specific acts as normatively "correct" or "incorrect". The most notable proponents of this idea are probably Haidt and colleagues (Haidt, 2001, 2004; Haidt & Joseph, 2004), who argue that all cultures have a strong normative preoccupation with matters of concern for others' suffering, social hierarchy, reciprocity towards others, and the taboo nature of certain proscribed actions, and that such concerns are therefore probably innate. Violations of norms concerning these three areas thus provoke strongly negative emotional responses amongst many of those who adhere to them.

It should, at this point, of course be noted that a large proportion of adults also frequently fails to behave normatively in many situations, and that adults too are prey to the same sort of self-serving biases as are children when it comes to behaving fairly. To use a well-known economic example, most "dictators" in the dictator game may offer something to the other player, but only a minority will offer a fair 50% split, even if this is often the modal offer (Engel, 2011). Whilst it is true that the majority of the population will cooperate conditionally in most situations, it is also the case that many who do so tend to "cheap-ride" by subtly undercutting slightly the amount of effort they are personally willing to put in relative to what they estimate the level will be in the rest of the "cooperating" population (Neugebauer, Perote, Schmidt, & Loos, 2009). Others are perfectly aware of social norms but discard them at will when they feel there is something to be gained by doing so.

Further complicating the issue is the fact that it is wholly possible for two internalised norms to give the actor conflicting advice. A classic economic example might be that of a responder in an ultimatum game who has just received a stingy offer. On the one hand, the norm of fairness has just been flagrantly violated by the proposer, which might mean that the norms of reciprocity and/or justice dictate that s/he should be punished by a rejection of the offer. On the other hand, there are also norms against doing harm to others and, given that the sums of money in a typical UG are generally not huge, against behaving in a way that could be construed as overly petty. A responder might equally decide to place more weight upon one of these and accept the offer, depending upon their personal preference. Finally, norms rarely function in a vacuum, and are often hard to isolate from other less pious motives. For example, rejection in a UG may reflect dissatisfaction at violation of the fairness norm, but less sophisticated factors such as inequity aversion and relative comparison may also be at work (Kurzban et al., 2015).

The exact norms which underlie the behaviour of subjects in behavioural economics can thus be somewhat opaque. Nevertheless, we can still infer a great deal from the individual's overt behaviour. Punitive behaviours signify negative other-regard, regardless of which actual norm helped to trigger the other-regarding concern behind the responder's actions. As discussed, the emergence of pro-social behaviour of a type that tends to be considered normative has been the focus of the vast majority of research on children's moral development to date. One of the questions posed by the present dissertation is whether greater evidence of punitive anti-sociality will also manifest itself to a similar degree throughout pre-adolescence (see, in particular, experimental Chapters 3, 4 and 5).

1.10: Punishment

Both fairness concerns and social norms, and the effect they can have on our other-regarding preferences, are vital to any discussion of hyper-competitiveness. The simple reason for this is that they act as two of the key motivators of punishment, which may be the raison d'etre for negative ORPs. Given that spiteful behaviour harms everybody involved, its existence is tricky to explain via natural selection, since behaviour which makes all parties worse off should not logically be selected for. The reasons for this existence almost certainly lie in humanity's unusually strong proclivity for punishment.

Punishment behaviour can take many forms. When an actor takes direct action against a target in response to an unwanted behaviour, this is known as second party punishment. This first, fairly basic, form of punishment can be "rational" in the economic sense, if it aims and succeeds in making the target desist from performing a harmful behaviour, to the actor's net benefit (Jensen, 2010; Jensen et al., 2007b). However, it can also violate the SEM. Humans will enact second party punishment against total strangers, even when there is no chance whatsoever of reaping any future benefit from doing so. Rejections in many one-shot ultimatum games are a classic example of such an act, and here we have a more likely candidate for an act motivated on some level by fairness concerns and norms. Certainly this type of act is "spiteful" at a behavioural level, since it imposes a net cost on all concerned (Levine, 1998). More sophisticated still is third party punishment (3PP), whereby an actor intervenes in a dispute between two others, even though s/he expends resources and has little to materially gain by doing so. This concern with a third party's outcomes is particularly important, since third party punishment experiments are a strong test of normativity in that they remove an actor's *material* self-interest as a plausible motivator of behaviour (Gummerum & Chu, 2014). Instead, the only remaining motivations are the sense of utility humans feel from observing that norms are adhered to, and norm violators punished. In economic terms, however, this behaviour is still spiteful, even if the benefit it conveys upon the victim of the original dispute is in some sense altruistic (Fehr & Gächter, 2002).

So, why does this colourful array of spiteful behaviour exist? Humans are a highly social species with a hugely complex network of interaction partners. This can provide useful material benefits in terms of reaping the fruits of collective or collaborative exercises, but it also leaves both the individual and the group vulnerable to the actions of "free-riders" – those who exploit the benefits of others' work without contributing to their production. Free-riding is classic

competitive behaviour, and is potentially very lucrative since it can provide gains whilst requiring the expending of little or no effort. It is therefore, perhaps unsurprisingly, a common human behaviour in all human cultural groups thus far tested (Chaudhuri, 2011; Henrich et al., 2006; Herrmann et al., 2008). It is also potentially a very damaging one since its presence has repeatedly been shown to cause group cooperation to cease, since most people have a tendency to "conditionally cooperate", meaning that they will do so only so long as they believe that others are doing likewise (Chaudhuri, 2011; Fehr & Gächter, 2002; Fischbacher, Gächter, & Fehr, 2001). However, if a facility exists whereby cooperators can punish free-riders, thus discouraging such exploitative behaviour, then there exists a mechanism to solve the free-rider problem and allow cooperation to remain stable. This theory has been successfully demonstrated empirically. Groups in economic games for which punishment of free-riders is an option achieve more stable cooperation than do those for whom it is not an option (Fehr, Fischbacher, & Gächter, 2002; Fehr & Gächter, 2002; Gächter, Renner, & Sefton, 2008; Gächter, 2014). Furthermore, given the choice between the two scenarios, subjects will preferentially choose a group in which sanctioning of free-riders is possible over one in which it is not possible, thus allowing the sanctioning group to dominate and out-perform the non-sanctioning group (Gürerk, Irlenbusch, & Rockenbach, 2006).

The importance of free-riding and punishment to human social interaction means that humans have both an exceptionally strong sensitivity to detecting when they are being cheated (Cosmides & Tooby, 1992, 2008), and a consistent willingness to punish and deter free-riders which appears to be ubiquitous to all cultures, in all of which it serves to effectively deter free-riders (Chaudhuri, 2011; Gächter et al., 2010; Henrich et al., 2006; Herrmann et al., 2008). Punishment, then, is central to human cooperation (Boyd & Richerson, 1992; Gächter, 2014). Obviously cooperation can also be reinforced by positive incentives (i.e. rewards), but experimental work has generally found punishment to be more efficient in most cases (Sasaki, Brännström, Dieckmann, & Sigmund, 2012), although the two mechanisms have also recently been observed to function particularly well in tandem (Andreoni, Harbaugh, & Vesterlund, 2002; Chen, Sasaki, Brännström, & Dieckmann, 2015). In the sphere of cooperative behaviour at least, it would appear that Machiavelli's maxim that it is best to be both loved and feared, but better to be feared if one cannot be both, holds true (Machiavelli, 1950). Particularly noteworthy in this respect is that in a long-term context with many interactions between actors, the simple *threat* of punishment is enough to ensure high levels of cooperation, meaning that the act itself rarely has to be administered so long as most subjects believe that there is a credible possibility that it might be (Gächter et al., 2008).

It is important to reiterate that costly punitive behaviour can violate the SEM just as systematically as does costly altruism. A responder's decision to decline any offer of greater than zero in an ultimatum game makes no more sense in rational choice terms than does the actor making an offer of greater than zero in the dictator game. It is because humans routinely engage in both types of act that they have been identified, uniquely, as "strong reciprocators" (Fehr & Fischbacher, 2003; Gintis, 2000b; Gintis, Bowles, Boyd, & Fehr, 2003), meaning that they will incur costs to pay back both generous and harmful behaviour in kind, even in cases where there is no chance of them reaping a net reward through doing so. From this perspective, negative ORPs such as envy and schadenfreude play a crucial, if somewhat counter-intuitive, role in supporting human cooperation, since they motivate punishment, the threat and enactment of which greatly deters free-riding behaviours which would otherwise cause cooperation to collapse.

This is not to say that negative ORPs are always a force for the social good. Indeed, in a subset of the population, spiteful costly behaviour appears to extend beyond mere punishment of the violation of the concept of fairness. Abbink & Herrmann (2011) showed that around 10% of subjects were prepared to pay a cost to ensure an initially even distribution became unevenly skewed in their favour, and that this rose to 25% when subjects were allowed to do so via a procedure which lent their actions a degree of plausible deniability. In other words, a substantial proportion of subjects were willing to pay an absolute cost to ensure relative advantage, notwithstanding the fact that most of them seemed to be aware that this was a socially undesirable behaviour, as witnessed by their greater reluctance to behave in this manner when they were unable to hide their actions. This "antisocial punishment" has been the focus of less empirical research than the functional punishment, outlined above, which serves to deter free-riders. It appears, at least in part, to be motivated by a particularly strong degree of social comparison via relative evaluation. This explanation alone cannot be sufficient, however, as results from public goods experiments have shown that a minority of subjects will pay a cost to spitefully deprive their most generous peers. Since these generous peers already have fewer resources due to their initial spending, the punishment aimed at them cannot simply be motivated by their greater relative wealth (Herrmann et al., 2008). Here, it would seem, is an example of SEM-violating behaviour in which subjects are willing to pay a cost not to eliminate inequity, but to actively increase it, albeit in their own favour.

Like most economic preferences, anti-social punishment is subject to considerable cross-cultural variation, but is present to some degree throughout the full range of cultures thus far

surveyed (Gächter et al., 2010; Herrmann et al., 2008) It is hard to explain via any of the motivators of ORPs referred to above, since it increases inequity (Fehr & Schmidt, 1999), damages social welfare (Charness & Rabin, 2002), violates the reciprocity norm and many others besides, it is unlikely to improve relative standing (Bolton & Ockenfels, 2000) compared to wider experimental sample as whole (only relative to the target specifically), and it harms those whose intentions were cooperative (Falk et al., 2003). Why it exists at all is currently uncertain, although one suggested theory is that it signifies a desire to accrue status through aggression and social dominance (Sylwester, Herrmann, & Bryson, 2013). While a proper understanding of anti-social punishment will not be possible until much more research has been carried out, one thing it does undoubtedly provide is a nice example of socially-motivated hyper-competitiveness in action. The prospect of gratuitously hyper-competitive behaviour undermining children's cooperation will be looked at in Chapter 6, and in particular Chapter 7.

Punishment is yet another common human social behaviour which has not been observed in our closest primate relatives, who have consistently failed to demonstrate it in a wide range of experimental contexts (Jensen et al., 2007a, 2007b; Kaiser et al., 2012; Proctor et al., 2013; Riedl et al., 2012). This represents a clear contrast with the actions of even pre-school children, who show an early proclivity for punitive behaviour. At as early as 3-years they appear to be capable of identifying shirkers and free-riders and excluding them from collective rewards (Melis, Altrichter, & Tomasello, 2013). More impressively, there exist several experiments demonstrating 3PP in pre-school children (Kenward & Östh, 2012; McAuliffe et al., 2015; Riedl et al., 2015; Robbins & Rochat, 2011; Salali, Juda, & Henrich, 2015), suggesting that the willingness to intervene on another's behalf for reasons that seem to be at least partly moralistic is a surprisingly early one to emerge ontogenetically. There also exists a comparative study which goes beyond this to show that 3PP does not occur in chimpanzees in directly analogous circumstances to those often tried with small children (Riedl et al., 2012), thus suggesting that 3PP may be yet another uniquely human lynchpin of the social order.

1.11: Focus of this dissertation

Given the importance of punishment to human cooperation, negative other-regarding concerns may be a vital part of human sociality. Relative to positive-other-regarding concerns, however, they have been the subject of scant focus. This is particularly true of the developmental literature, which means we know very little about how they emerge and

develop. Whilst there is an increasing focus on their presence in the pre-school years and even infancy, for example in the recent spate of 3PP studies cited in the previous section, there is much less in the way of comparative work which charts punishment behaviours through childhood to adolescence. There is also very little evidence on how punishment actually affects children's behaviour (although see Lergetporer et al., 2014). Furthermore, hyper-competitive behaviour, and the negative ORPs that underpin it, has tended to be looked at in isolation. Whilst its existence is established, we know little about how it compares to, and interacts with, hyper-cooperative behaviour, or how it affects children's abilities to make cooperative decisions.

In this dissertation I have adapted a series of eight experiments from behavioural economics and game theory in an effort to directly explore negative ORPs in children, and in doing so contribute to the rectification of this imbalance. My aim is not to provide a new, over-arching theory of negative ORPs, but rather to provide a wealth of new empirical data from which future experimental and theoretical work can draw. Based on the preceding literature review, I have identified the following areas in particular for scrutiny:

1) How does the strength of negative ORPs change across the breadth of childhood? Does it increase or decrease? Is this process linear, and does it occur reliably across different experiments?
2) How do negative ORPs compare in this respect to positive ORPs? Do they develop in the same way at the same time and at the same pace? Can the two be directly compared? Is a tendency to violate the SEM something we simply learn with age, and are all violations of the SEM driven by the same preoccupation with fairness?
3) What can the expression of ORPs in childhood tell us about the developing understanding of fairness preferences and norms?
4) What are the implications of negative ORPs for children's ability to cooperate? Negative ORPs tend to be looked at in isolation, i.e. in experiments where punishment is the focus and to punish or not to punish is the outcome variable. How do negative ORPs affect children's behaviours in other contexts?

These are broad questions, but the eight experiments contained in this dissertation aim, if not to answer them then to explore them, and in doing so to pave the way for future empirical and theoretical work. All eight of these experiments were conducted in a specific and slightly unusual setting, namely the Science Museum in London. Collecting data there allowed me to test a very large volume of subjects – data from over 1,500 of them is included in this

dissertation, making it one of the largest, if not the largest, in depth studies of negative ORPs in children ever performed. Nevertheless, there are certain unusual features of collecting data in such an environment, and Chapter 2 will explain these, as well as setting out some of the other methodological considerations which were common to all of the experimental studies. Doing so will give the reader a better sense of context and also save repetition and cross-referencing in later chapters.

In Chapter 3 I look directly at spiteful, second party punishment in children using a specially adapted version of the ultimatum game. The subjects tested range in age from pre-schoolers to the early stage of adolescence. The aim of this experiment was to definitively establish the age at which children routinely express a preference for punitive behaviour. The focus of Chapter 4 is similar, featuring a variant of the same experiment, but it aims to extend this discussion by focusing in more detail on the distinction between outcome and intention, and whether children of various ages understand it. It also successfully replicates some of the more unusual findings witnessed in Chapter 3.

Chapter 5 extends the study of punishment, or negative reciprocity, by placing it alongside mutual rewarding, or positive reciprocity, in a moonlighting game, in an effort to assess the strength of both tendencies. Few experiments with either adults or children have ever examined both types of reciprocity simultaneously in a directly comparable way, yet this is a very important comparison. Conclusions about our pro- or anti-sociality as a species can be far better informed by studying the two tendencies in tandem rather than individually.

Chapter 6 seeks to extend the scope of the dissertation somewhat by looking at hyper-competitive behaviour in a broader context. Chapters 3-5 primarily focus on costly punishment. Chapter 6, however, seeks to explore how hyper-competitive tendencies affect cooperation. It does this by placing children in a dynamic, strategic scenario, the battle of the sexes game, and posing the question of if and how their known tendency to think in zero-sum terms will interfere with cooperative, pro-social outcomes.

Chapter 7 also focuses on hyper-competitive motivations in a strategic context, but this time the context is more unambiguously cooperative. The stag hunt is the classic model of mutualistic cooperation, which should be the simplest form of cooperation to induce. If hyper-competitive motivations are observed to hamstring cooperation in such an environment, this will have very interesting implications for human sociality. Evidence from the chicken game, effectively a reverse stag hunt, is also considered.

Chapter 8 will conclude the thesis with a general discussion of the implications of the experimental chapters for our hyper-competiveness, and therefore for our wider social behaviour generally.

Chapter 2: Data Collection Methods and Considerations

2.1: The Live Science Area

The data presented in all five of the following experimental chapters were collected at the Wellcome Trust "Live Science Area" at the Science Museum, London. The Live Science Area is an interactive research exhibit in one of the Museum's galleries, and is accessible to all visiting members of the public during regular opening hours. The Area is offered to teams of visiting researchers for rotating residencies of three months, during which a single research team has sole access to the Area's facilities and data collection opportunities. The data presented in this dissertation were collected during two separate residencies, with assistance from a team of student interns. The data presented in Chapters 3 and 4 were collected between October and December 2012. The data presented in Chapters 5-7 were collected between January and March 2014.

2.2: Sampling

Due to the nature of the Live Science Area, recruitment for all of the experiments was done by opportunity sampling. Researchers would be based in the exhibit and would greet and test volunteering members of the public on an ad hoc basis. All experiments were approved by the Research Ethics Committee at Queen Mary University of London, and all subjects were not permitted to begin testing until the research team had received informed, written consent from their parents.

Via this method a very large volume of data was able to be collected in a comparatively short period of time. Data from over 1,500 subjects are presented in this dissertation. This notwithstanding, some additional considerations caused by this set-up which are different to conventional sampling methods need to be taken into account. Chief amongst these is that because the recruitment method was spontaneous and ad hoc, I could not control as perfectly for the demographic composition of the sample as is often the case in conventional developmental experiments, in which subjects are individually selected to take part from an existing database, based upon their demographics. For example, we were commonly presented with pairs of siblings who wished to participate. Testing siblings together is an

unusual feature of a developmental psychology experiment, but they were nonetheless accepted for all of the experiments described here. This was done partly because part of the remit of visiting researchers at the Science Museum is to provide an inclusive service to all of the Museum's visitors, and partly in the interests of collecting as large and diverse a sample as possible.

Nevertheless, this demographic diversity had the potential to cause considerable "noise" in the results. Consequently, in all experiments, the following three factors are included in the analysis. The first is participants' age in months. Given that the focus of this dissertation is the emergence and development of other-regarding preferences, this was a variable of central interest. Consequently, the influence of age (or lack thereof) on the key outcome variables, such as levels of reciprocity and willingness to cooperate, will be reported in all experimental chapters. The second demographic factor is participants' gender. Although the issue of whether hyper-competitiveness varies in boys and girls is un-explored and, potentially, of considerable theoretical interest, it was not the primary focus of this dissertation. Nevertheless, previous work on children's social behaviour found that the genders sometimes differ in their degree of prosociality and that, when they do, this difference tends to manifest itself as greater prosociality in girls (Eisenberg, Fabes, & Spinrad, 1998). As a result, subjects' gender is reported as a factor in all data analysis, in order to control for this confound if and when it occurs.

The third demographic factor is participants' degree of pre-acquaintance with one another, i.e. whether or not they had met prior to the experiment beginning, and in what capacity. Unlike with gender, it is less easy to make predictions based on previous research as to how this factor should affect results, since testing friends and siblings together is relatively unusual within developmental psychology. The previous work on the topic which has been done suggests than in pre-school and middle childhood children will sometimes distinguish between friends and strangers when making economic decisions (C. Moore, 2009; Olson & Spelke, 2008). There is also evidence that, starting at about 5-6 years and increasing in importance thereafter in conjunction with increasing age, children begin to pay growing attention to whether their interaction partners are ingroup or outgroup members, discriminating ever more systematically in favour of the former (Angerer et al., 2015; Buttelmann & Böhm, 2014; Fehr et al., 2013; Gummerum et al., 2009). Although this area was, again, not my primary focus, such existing results necessitated including whether or not children were pre-acquainted with their partner as a between-subjects factor in the analysis.

All of the experiments reported in this dissertation involved children being tested in pairs. In Chapters 3-5, the two children within these pairs had different roles, a "proposer" and a "responder". Consequently, in the analysis of these experiments, these different roles are examined separately, with the individual child the focus of the analysis. In all cases, subjects' age in months, gender and relationship with partner (known/unknown) are included as between-subjects factors.

In Chapters 6-7, the two children within each pair had essentially the same role, with identical and simultaneous choices presented to both partners on each trial, and the end result of each trial the joint product of both partners' decision. These experiments also featured a large number of repeated trials, each consisting of 20-30 iterations. In these experiments, subjects' patterns of behaviour were almost certainly systematically affected by the past experiences of joint outcomes, meaning that truly independent observations at the individual level are not possible. Given this, the pair ("dyad") and not the individual is taken as the primary unit of analysis. Dyads' mean age in months, gender composition (all-male/all-female/mixed) and relationship (known/unknown) are included as between-subjects factors. The one exception to this rule is data recorded on the first trial of each experiment, at which point subjects' behaviours had not yet had a chance to be influenced by their partners' choices. On these rounds only, therefore, data from individuals were also considered, using the same demographic predictors as in Chapters 3-5.

2.3: Coding and Data Analysis

During all of the experiments, data were live coded during the procedure by the experimenter. All trials in all experiments were recorded on video camera for coding purposes and then re-watched and cross-checked in full back at the lab, before data entry and analysis were begun. All subjects and their parents were fully aware of this, and the written consent forms made explicit reference to the fact that video coding would be part of the research process.

Specific outcome variables in each experiment will be explained in depth in the relevant chapters. Generally, however, the data consisted of observing and tallying simple binary choices between a more cooperative and a less cooperative action, indicated either by pulling a lever or playing a card. The only exception to this is in Chapter 5 (the moonlighting game), in which the outcome variable on each trial was derived from adding small numbers of points (from 0-8). Where possible, parametric statistics are used in this dissertation. However, there

are two scenarios in which they are not used. The first is when the principles of homogeneity of variance and/or sphericity are violated, in which case equivalent non-parametric tests are used (e.g. Friedman's test instead of ANOVA in Chapter 6). The second is when binary outcomes of individual trials are analysed, in which case methods more appropriate to binary data are used, such as logistic regression, Cochran's test, McNemar's test, chi-square analysis and binomial analysis (Field, 2009).

2.4: Testing conditions at the Live Science Area

The Live Science Area at the Science Museum is a physical installation. Consequently, it has certain un-alterable parameters which had to be considered when designing and running the experiments. As with the sampling issues described above, this meant that certain procedural features of the experiments were unusual when compared with standard procedures in the wider field. The first is that the Area is partitioned from the rest of the gallery on three sides, but not totally enclosed. This means that watching members of the public could view the experiments in progress, albeit from a distance and from out of participants' line of sight. Parents and family members were also invited to watch proceedings take place as a feature of the exhibit. This, of course, means that the trials were not carried out anonymously, and subjects were aware of this and the fact that their decisions were therefore potentially open to a degree of public scrutiny. There is some existing evidence to suggest that children tend to adjust their behaviour to make it more socially desirable when watched (Engelmann et al., 2012; Engelmann, Over, Herrmann, & Tomasello, 2013; Leimgruber et al., 2012; A. Shaw et al., 2014). This therefore needs to be considered when discussing the overall findings.

Also important to note is that this set-up also meant that, of course, subjects' decisions were visible to each other and to the experimenter. When behavioural economics experiments are performed with adults, this is almost never the case, for two reasons. First, subjects are almost always significantly more generous when they know their behaviour is being observed, seemingly due to social desirability concerns and a desire to manage their reputations (Andreoni & Petrie, 2004; Hardy & Van Vugt, 2006; Piazza & Bering, 2008; Satow, 1975). Second, the more personal information subjects have about their fellow participant, the more likely they seemingly are to make prosocial decisions (Andreoni & Petrie, 2004; Lamba & Mace, 2010; Rege & Telle, 2004). In developmental psychology, this constraint is sometimes less rigorously applied, and it is not difficult to think of multiple examples of successful studies in which children have been tested together in the manner of this dissertation (Blake &

McAuliffe, 2011; Fan, 2000; House et al., 2013; Leimgruber et al., 2012; McAuliffe et al., 2013; Takagishi, Kameshima, Schug, Koizumi, & Yamagishi, 2010; Vogelsang, Jensen, Kirschner, Tennie, & Tomasello, 2014; Wittig, Jensen, & Tomasello, 2013). Again, however, the potential impact of the public nature of decision making must be borne in mind.

The above work on observer effects notwithstanding, I chose to keep the process of decision making comparatively public for two reasons. The first was to ensure maximum transparency, and remove any possible doubt about which agent was causing the economic decisions on display. The second was in the hope of allowing overt social comparison, in the interests of thus also allowing potential hyper-competitiveness to arise.

2.5: Points, Prizes and Rewards – the Scoring System

A final important methodological feature common to all of these experiments is the way in which participants were incentivised to take part. In all of the experimental chapters, participants' decisions on each experimental trial scored them "points", to be exchanged for prizes once all trials had been completed. This use of points as a pseudo-currency has been successfully employed in many previous developmental experiments (Fan, 2000; Harbaugh & Krause, 2000; House et al., 2013; McCrink, Bloom, & Santos, 2010; Robbins & Rochat, 2011; Sheskin, Bloom, et al., 2014; Steinbeis & Singer, 2013). In the present series of experiments, points consisted of plastic gumball capsules, which were physically awarded to participants in the appropriate number after each trial, by being placed in a transparent plastic bucket adjacent to where they were sat.

All of the experiments involved multiple rounds with multiple opportunities to accrue points. Crucially, as both participants' prize buckets were transparent, subjects had a clear and easy means of comparing their allocation to that of their dyadic partner's at all times. The decision to allow participants to watch both their own rewards and those of the other player physically accumulating is an unusual novelty even in the developmental literature. For example, in a previous study which looked at the impact of public and private rewards upon children's prosociality, points in all conditions were removed from sight and hidden in a felt bag at the end of each trial (Leimgruber et al., 2012).

My methodology deliberately employed completely the opposite approach, and I feel there is sound justification for this. Conventionally, developmental economics experiments remove all

accrued rewards from participants' field of vision after each round. However, unlike adult participants, who in repeated-rounds experiments are generally given explicit visual feedback as to what their overall score is, and sometimes even how it compares to that of other players (typically on a computer screen), this means that the children are often left with little to no information about how they are actually doing. Their endowments simply appear by magic and then disappear just as mysteriously from sight and mind at the end of each experimental trial. It is thus not entirely clear what the children, particularly at younger ages, know or remember about their cumulative score and how it compares to that of other players. This has serious implications for their ability to compare their own score to that of others, and thus draw relative, other-regarding comparisons. In all experiments in the present dissertation, they were not left in any doubt as to how they or the other player were doing. Whether or not this succeeded in triggering hyper-competitive behaviour will be a matter of considerable interest.

In all experiments, points were exchangeable in units of ten for small gifts and prizes from the Science Museum's gift shop. For every ten points a child accrued, they could purchase an extra prize, with the bigger and better prizes only available to those with high scores. So, for example, ten points would entitle them to a badge, twenty points to a badge and a pencil, thirty points to a badge, a pencil and a bouncy ball, etc. This scoring system was always explained in full to participants before each experiment began, and participants were reminded several times that the best way to get prizes was to "fill up your buckets as fast as you can".

2.6: Conclusion

Collectively, this dissertation consists of the largest study of negative other-regarding preferences in childhood ever conducted. Although data collection in public places imposes certain restrictions, I feel these were more than compensated for by the opportunity to test a large and relatively diverse sample, drawn from across an unusually large age spectrum, in multiple experimental paradigms. Any limitations were mitigated by careful experimental design and appropriate inclusion of potential confounds in the analysis. The large sample sizes and relative consistency of subjects' behaviours across studies allows for confidence in the findings presented in the next five Chapters.

Chapter 3: The mini-ultimatum game with theft

3.1: Introduction

As Chapter 1 made clear, the ultimatum game (UG; Guth et al., 1982) is the classic economic experiment for examining negative other-regard. When a responder rejects a proposer's offer, they are paying a concrete cost to punish what they perceive as unfair behaviour targeted towards themselves. The mini-ultimatum game (MUG; Falk, Fehr & Fischbacher, 2003) builds on this basic scenario by presenting responders with a series of offers in which unfairness is sometimes the product of the proposer's stinginess, but other times of the experimenter's design. In this way it allows a shrewd responder to differentiate the extent to which small offers are genuinely the products of meanness on the part of the proposer. This chapter will first offer a quick overview of the growing literature on UGs and MUGs in children. It will then introduce the mini-ultimatum game with theft (Kaiser et al., 2012), a novel experimental extension of the MUG which increases the emphasis on the proposer's intentional role in creating unfair outcomes. This novelty should make proposer meanness particularly flagrant, and thus be well-suited to motivating negative other-regarding concerns amongst responders at even the youngest ages tested. The age range of participants in the present experiment is from 4-13 years. In conjunction with a replication and modification of the MUG with theft (MUGWT) in Chapter 4, it will also constitute the largest UG study of any kind yet performed with a developmental sample.

3.1.1: Aims and Objectives

The aims and objectives of Chapter 3 are as follows:

1) To test whether or not pre-school children will violate the standard economic model (SEM) as responders in the UG by enacting second-party punishment.
2) To investigate whether preferences for second party punishment become monotonically more common with age, pre-school onwards.
3) To examine UG behaviour throughout childhood, from pre-school to adolescence, using a single methodology, in order to examine how other-regarding preferences change at different stages of development.
4) To test fairness preferences throughout this age spectrum, by looking at changes in proposer behaviour.

5) To adapt and apply the mini-ultimatum game with theft to human subjects for the first time.

3.1.2: Children's Preferences in the Ultimatum Game and Mini-Ultimatum Game

Children's Preferences as Responders

When examining negative other-regarding preferences using UG experiments, the responder's role is the one of primary interest. This is because it is the responder's rejections which violate the SEM for spiteful reasons. By rejecting low offers, the responder pays a cost to ensure that the proposer loses their allocation, an act which is essentially a form of second-party punishment.

There are around a dozen previous UG-based studies with children as subjects. Children's preferences as responders have thus far not been entirely consistent. In the majority of studies using a conventional UG (i.e. one in which the proposer's offers are chosen from along a continuum beginning at zero), punitive violations of the SEM have tended to increase monotonically in frequency from mid to late childhood. In other words the older children get, the more likely they are to deviate from the SEM by rejecting offers to punish meanness (Harbaugh, Krause, & Liday, 2003; Sally & Hill, 2006; Steinbeis et al., 2012). This pattern is also replicated in a developmental MUG focusing on those at 5, 8 and 12-years (Bereby-Meyer & Fiks, 2013).

In contrast, however, the first developmental UG did not find any clear relationship between age and the frequency of responder rejections (Murnighan & Saxon, 1998). This is important since not only was this a large study, it was also one of very few to look at a broad spectrum of development, involving subjects from pre-school to adolescence. Furthermore, most of the MUGs performed to date have typically not found any linear increases in costly rejections of unfair offers with age (Gummerum & Chu, 2014; Güroğlu, van den Bos, & Crone, 2009; Sutter, 2007), although Güroğlu et al. (2009) did observe a lesser likelihood amongst 9-year olds to punish offers of 25% than was seen amongst participants aged 12 and above. Unlike Murnighan and Saxon (1998), however, most of these studies were focused on children from mid-to-late childhood, with no data collected on pre-schoolers. Pre-schoolers, as noted in Chapter 1, are a particularly important source of data on other-regarding preferences, due to their lesser normative internalisation and perspective-taking abilities.

Whilst the data in these studies may not be entirely complementary, they do contain certain consistencies. For example, all responders tested aged 6-years and upwards in the studies reviewed so far have been willing to violate the SEM (Bereby-Meyer & Fiks, 2013; Gummerum & Chu, 2014; Güroğlu et al., 2009; Harbaugh et al., 2003; Murnighan & Saxon, 1998; Sally & Hill, 2006; Steinbeis et al., 2012; Sutter & Kocher, 2007). Furthermore, if and when age effects have occurred, they have tended to show either an increase in second party punishment with age or a non-linear age pattern which nevertheless features those in late childhood as amongst the most punitive – a *reverse* linear age effect whereby a willingness to violate the SEM decreases has never been observed. The cumulative pattern therefore seems to be for a slow increase in other-regard with age, albeit one which does not necessarily follow a linear pattern.

This raises the question of when in ontogeny this tendency begins, but so far this remains relatively unexplored. There is little data on pre-schoolers' preferences in the UG, and that which does exist does not present a clear picture. Put simply, it remains to be conclusively established whether pre-schoolers have sufficient negative other-regarding preferences to violate the SEM. For example, Bereby-Meyer and Fiks (2012) found that 5-year olds' behaviour does not differ from the SEM, in that they reject offers of zero, but nothing greater than zero. Other studies have observed a significant, albeit small, proportion of those aged 5-years and under willing to make costly rejections of stingy offers (Murnighan & Saxon, 1998; Takagishi et al., 2014; Wittig et al., 2013). By contrast, Takagishi et al. (2010) found such behaviour to be present in a majority of subjects of this age, although it should be acknowledged that their sample numbered only 11.

In addition to these inconsistencies as to what pre-schoolers preferences actually are, there are very few experiments which have attempted to directly compare them to those of older children using a single methodology, thus rendering direct comparisons difficult. I know of only two, and these unearthed largely contradictory findings. Bereby-Meyer & Fiks (2013) observed that pre-schoolers adhere to the SEM, and that this tendency disappears by 8-years and continues to diminish into adolescence. Murnighan and Saxon (1998), however, found that 5-year olds violated the SEM, and that the tendency to do so showed no linear age effects from thereon upwards.

Children's Preferences as Proposers

The proposer's role in the UG is of secondary interest for the purposes of this dissertation. Although the proposer too shows increasing deviation from the SEM the more generous their offer is, this is of less direct relevance to a study of hyper-competitiveness. Firstly, proposer offers are evidence of positive rather than negative other-regarding preferences (ORPs), and secondly it is difficult to disentangle whether generous proposer behaviour is attributable to a genuine concern with fairness norms, or simply a strategic desire to avoid punishment (Sally & Hill, 2006), although it is true that in either case the proposer's generosity results from having taken the responder's perspective into account.

For proposers, too, there are inconsistencies in previous findings which would benefit from further exploration, and which the present experiment can contribute to resolving. Proposer offer sizes are sometimes observed to be larger in elder age groups (Harbaugh et al., 2003; Kogut, 2012; Steinbeis et al., 2012), although here there also exist examples of studies in which no linear effect of age on deviation from the SEM is present (Murnighan & Saxon, 1998; Sally & Hill, 2006; Sutter, 2007).

Once again, there are relatively few existing studies on pre-schoolers and even fewer which compare their preferences to those of older children. Of those which exist, one UG study found pre-schoolers to be relatively generous (Murnighan & Saxon, 1998), with two MUGs finding similarly (Sally & Hill, 2006; Takagishi et al., 2010), albeit with the caveat in the case of the latter that this was only true of those with higher-functioning theory of mind abilities. By contrast, however, two more recent UG studies found pre-schoolers to be extremely ungenerous, sometimes barely differing from the SEM at all (Takagishi et al., 2014; Wittig et al., 2013). Only three studies have attempted to compare pre-school proposers in the UG to older children. Of these, one found them to be significantly less generous (Kogut, 2012), whilst the others found no linear age increase from pre-school to later childhood (Murnighan & Saxon, 1998; Sally & Hill, 2006).

As proposers too, then, children's fairness preferences remain rather opaque, although it is true that few if any authors have suggested that they are *more* likely to violate the SEM with generous offers than are older children. In addition, conceptually similar experiments comparing young children's distributional preferences generally find them to be largely self-regarding and ungenerous prior to the age of 8-years (Bügelmayer & Spiess, 2014; Fehr et al., 2008; House et al., 2013; Sheskin, Bloom, et al., 2014; Smith et al., 2013). As a consequence, I

will base my assumptions in the present experiment upon these wider findings and predict less generous behaviour amongst younger proposers, even if the previous evidence from the M/UG specifically offers only partial support for this.

3.1.3: Mini-Ultimatum Game with Theft

Previous work is unclear as to whether pre-school children are willing to enact spiteful second-party punishment. Whether and to what extent they are willing to do this is a key focus of this Chapter. I therefore implemented a methodology designed to make the proposer's role in causing inequity and unfairness particularly salient to the responder. The "mini-ultimatum game with theft" (MUGWT; Kaiser et al, 2012) has never before been applied to human subjects. Its only previous incarnation has been with chimpanzees and bonobos, species which UG research consistently finds unable to deviate from the SEM and reject low offers in ultimatum games (Jensen et al., 2007a; Kaiser et al., 2012; Proctor et al., 2013). Comparing the results in this chapter to the primate responders in Kaiser et al. (2012) will thus serve as a useful guide as to whether negative ORPs were observed in the present experiment.

The MUGWT is similar to the MUG in that the proposer makes a binary choice between a comparatively generous offer and a comparatively ungenerous one. However, in the MUGWT the ungenerous offer is made by "theft". In practice this means that some or all of the initial allocation is placed in front of the responder, and then subtracted by the proposer. This is in contrast to a regular UG/MUG, in which typically the proposer begins with the entire allocation, and is left to decide how much they wish to offer. In other words, the UG's implicit giving framework is replaced with one which emphasises taking. Spontaneous giving of coveted resources is not naturalistic behaviour amongst non-human primates, and the same appears to be true of small children (Birch & Billman, 1986; LoBue et al., 2011; Svetlova, Nichols, & Brownell, 2010). However, from around 3-5 years of age upwards Western children do appear to understand and value the sanctity of others' property, and will protest at the violation of this principle (Kanngiesser et al., 2015; Rossano et al., 2011). Theft, therefore, ought to be recognised as such, and to provoke considerable objection even amongst the youngest of the subjects in the present experiment.

Children's behaviour as proposers in MUGWT will therefore be of interest, particularly in the case of those below 7-8 years, who may be likely to be willing to discard the fairness norm (Smith et al., 2013). In the primate version of the MUGWT (Kaiser et al., 2012), proposers, like

responders, were wholly self-regarding and chose the maximum amount for themselves on every trial. Human adults in the regular MUG behave very differently, for example they almost never allocate 100% of the endowment to themselves (Falk et al., 2003).

Wholly self-regarding behaviour also seems unlikely from most of the children in our sample (4-13 years), but the very youngest proposers may possibly be an exception to this rule. Wittig et al (2013) found 5-year proposers to be both selfish in their allocations and poor at anticipating offer rejection in a recent MUG, although it is true that even they were less overwhelmingly greedy than chimpanzees and bonobos in both a MUG and a MUGWT (Jensen et al., 2007a; Kaiser et al., 2012). This notwithstanding, the children's behaviour in Wittig et al (2013) was generally not consistent with a serious concern with fairness. When combined with the fact that many young children have limited inhibitory control (Steinbeis et al., 2012) it may be that the "theft" framing of the MUGWT is too tempting for proposers to resist. On the other hand, pre-school children have been observed making spontaneous offers in the dictator game (Benenson, Pascoe, & Radmore, 2007; Gummerum, Hanoch, Keller, Parsons, & Hummel, 2010), and even at 4-5 years their perspective-taking abilities are beginning to eclipse, or are at the very least achieve parity with, those of primates (Call & Tomasello, 2008). Whether this is sufficient to allow them to inhibit theft in early childhood will be interesting to see. Whatever the outcomes at that age, by around 8-years of age it is to be expected that proposer behaviour will be significantly more egalitarian, based on the big strides in this area witnessed at around this age in many previous studies (Blake & McAuliffe, 2011; Fehr et al., 2008; House et al., 2013; Kogut, 2012; McAuliffe et al., 2013; McAuliffe et al., 2014; Sheskin, Bloom, et al., 2014; Smith et al., 2013).

3.1.4: Hypotheses

There should be much the MUGWT can tell us about children's negative other regarding preferences. It can also help to expand the literature on children's preferences as proposers in the UG. In both respects, the recruitment of subjects from across a wide age spectrum should be of particular benefit. Based on previous findings to date, this chapter will test the following hypotheses:

- **(H1)** All responders, including those in their pre-school years, will violate the SEM by displaying costly, second-party punishment.

- **(H2)** The frequency with which children prefer costly second-party punishment will increase monotonically with age. Older children will be more willing to violate the SEM.
- **(H3)** Responders across the age spectrum will be more consistent in their preferences towards other kinds of offers. Acceptance rates of fair and advantageous offers are expected to be consistently high. Cost-free punishment (i.e. rejection) of zero offers is also expected to be consistently high.
- **(H4)** Proposer offers will become more generous with age in conditions where theft causes advantageous inequity. In particular, children aged 8-years and upwards should be more willing to make fair offers than are younger children in these conditions.

3.2: Methods

3.2.1: Participants

All subjects were recruited at the Live Science Area in the manner described in Chapter 2. Issues of sampling and informed consent were thus as explained in the previous chapter. Children were tested face to face in pairs. All decisions taken by both proposer and responder were their own, and no deception was involved at any point.

Three hundred children aged 4-13 years took part in the study. They were tested in pairs, within which one subject was randomly allocated to the role of the proposer, and the other to the role of responder. Of the 150 proposers, 82 were girls and 68 were boys, with an average age of 8.75 years (SD=2.45 years). Of the 150 responders, there were 64 girls and 86 boys, with an average age of 8.25 years (SD=2.54). During analysis the sample was broken down into four categorical age cohorts. The cohorts represent the stages of pre-school (4-5 years), pre-fairness preferences (6-7 years), emergent fairness preferences (8-9 years) and post-fairness preferences (10+ years), based on findings from previous studies (e.g. Blake & McAuliffe 2011; Fehr et al, 2008; House et al 2013a). Within the proposer dataset the numbers within each of these cohorts were as follows: 4-5 years, n=19; 6-7 years, n=43; 8-9 years, n=40; 10+ years, n=48. For the responder dataset the breakdown was: 4-5 years, n=30; 6-7 years, n=44; 8-9 years, n=32; 10+ years, n=44.

3.2.2: Apparatus

The experimental set-up consisted of an apparatus placed on the floor between two chairs. The apparatus is shown in Figure 1. One child would be sat to the left of the apparatus (the proposer) and the other to the right (the responder). The apparatus consisted of a wooden box of 80cm in length, 35cm in width and 40cm in height. The box was designed with three handles, two on the proposer's side and one on the responder's. The upper surface of the box rotated on two axes; vertically towards the proposer, an action cause by the proposer's theft handle, and horizontally, an action caused by the responder's handle. Points were distributed at either end of the apparatus according to experimental condition (see Design, section 3.2.3), and were moved around by manipulation of the three handles.

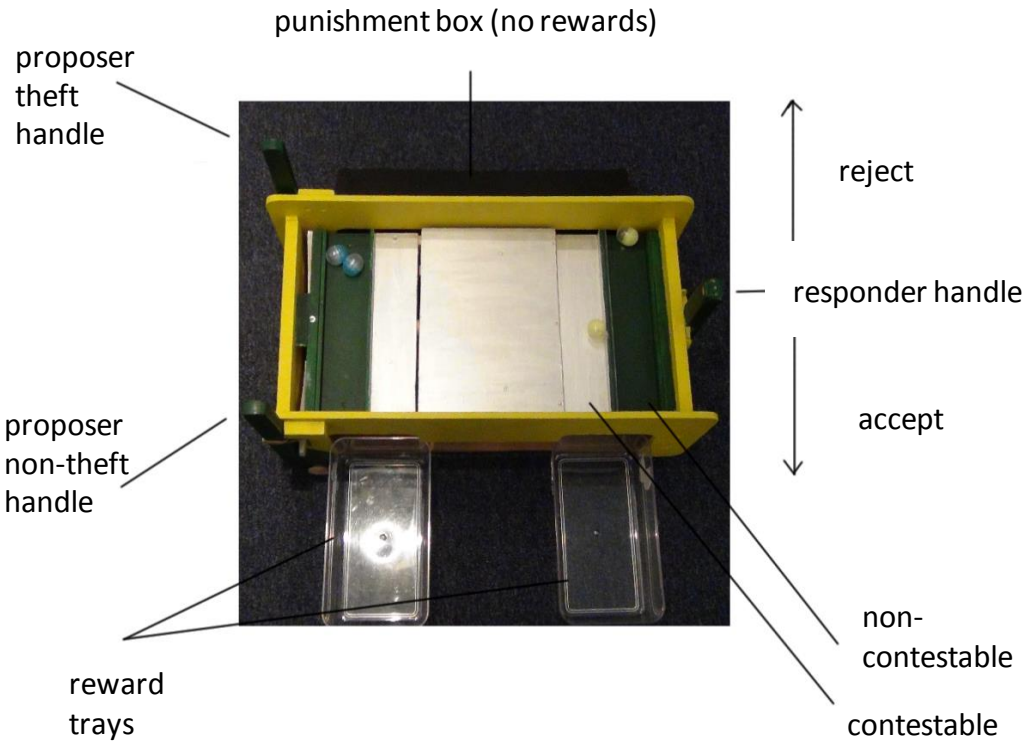

Figure 1: Apparatus used in Experiments 1 and 2

This is a bird's eye view of the apparatus used in Experiments 1 and 2, using the selfish theft (2/1;1) condition as an example. In the default set-up shown, both participants start with two points adjacent to their reward tray. The proposer moves first and has the choice of whether to steal the reward ball in the responder's contestable area via a tipping mechanism activated by the theft handle, or to leave the rewards as they are via the non-theft handle. The responder then decides to accept or reject the resulting distribution by tipping their handle left or right accordingly.

3.2.3: Design

The design had both between and within-subjects elements. Subjects only performed one role for the duration of the experiment, i.e. every participant was *either* proposer or responder, but never both. However, all of these pairs were presented with seven trials, each consisting of one of the seven experimental conditions shown in Table 2.

In all conditions, four points were divided between three different areas of the playing surface, and the notation used to explain each condition in this Chapter and in Chapter 4 reflects this. In each case, the number on the left shows the proposer's default allocation, the number on the right shows the part of responder's default allocation which is *safe from theft*, and the number in the middle shows the part of the responder's default allocation which *can be stolen*, i.e. transferred from responder to proposer. So, for example, in the "greedy theft" condition, the notation used is "2/2;0". What this means is that the proposer's default allocation is 2, the responder's *safe* default allocation is 0, and the responder's default allocation which *can be stolen* is 2. In other words, both participants start with an allocation of 2, but in this condition both of the responder's tokens can be stolen by the proposer. The proposer can therefore make a fair and equal offer by not stealing (leaving the default allocation of two each), or a very unfair offer via stealing, which results in everything for his or herself and zero for the responder.

Table 2: the seven conditions in Experiment 1.

condition	*notation*	*initial distribution*	*distribution after theft*
1. greedy theft	2/2;0	2 - 2	4 - 0
2. fair theft	1/1;2	1 - 3	2 - 2
3. pointless theft	2/0;2	2 - 2	2 - 2
4. selfish theft	2/1;1	2 - 2	3 - 1
5. modest theft	0/1;3	0 - 4	1 - 3
6. responder advantage	0/0;4	0 - 4	0 - 4
7. zero-sum	0/4;0	0 - 4	4 - 0

As reflected by the names in Table 2, the differing proposer offers in each condition represent different levels of provocation and unfairness, from the justifiable (fair theft, which leaves both subjects with two points) to the outrageous (greedy theft, described above, which leaves the proposer with 100% and the responder with zero). Conditions 1-4 were based on those of Kaiser et al, 2012, but are further simplified for the benefit of the youngest participants in that the number of points in play within a single trial was reduced from ten to four. Conditions 5-7

are novel additions to the MUGWT intended as additional controls. In conditions 5 (modest theft control) and 7 (zero-sum control) a rational proposer should always steal, since they start with nothing, though the rejection rates between the two should be radically different given that the proposer leaves the responder with either 75% (modest theft) or nothing (zero-sum). In condition 6 (responder advantage control) a rational responder should never reject regardless of the proposer's choice, since the responder is guaranteed 100% of the allocation in either case.

3.2.4: Procedure

The total procedure took approximately 15 minutes and was split between a familiarisation phase of approximately 8-10 minutes, in which subjects were shown how to work the apparatus, and then a testing phase of around 5-7 minutes in which they experienced all seven of the conditions displayed in Table 2 in a counterbalanced, pseudo-random order.

It is important to note that care was taken throughout all stages of the experiment to avoid priming children with language suggestive of either cooperation or competition. The procedure was referred to as an "experiment" and not a "game", and all terms like "fair", "steal" and "punish" were omitted. For example, the Theft handle was referred to as the "tipping" handle, which is a neutral description of its physical action.

Familiarisation phase

The proposer and responder were sat at opposite ends of the box. The proposer's end had two handles, one for making an offer by stealing (the Theft handle) and one for making an offer by accepting the default allocation (the Release handle). The latter was so called because the only effect of pulling on it was to open a locking mechanism in the apparatus, which allowed the responder to make a move. The Theft handle, by contrast, both opened the locking mechanism *and* stole any points placed in the responder's contestable area (see Figure 1). The theft was made by the Theft handle tipping the playing surface vertically towards the proposer, thus causing the points in the responder's contestable area to roll down and onto the proposer's end of the playing surface. Importantly, any points in the responder's non-contestable area were prevented from rolling across during the tipping action by a Plexiglas barrier.

Once the proposer had made a choice and pulled one of their handles, it became the responder's turn to move. The responder had one handle, which could rotate horizontally to either the left or the right. If the responder wished to *accept* the proposer's offer, then tipping the responder handle to the left caused the points to roll into two transparent reward trays, one at the proposer's end of the apparatus and one at the responder's. All points in these trays counted towards subjects' final scores. If the responder wished to *reject* the proposer's offer then tipping the responder handle to the right caused the points at each end to roll into "the black box". Any points tipped into here were eliminated from the game and given to no-one (i.e. this was the responder's spiteful punishment option).

The *responder's* role was explained first. Pilot testing revealed that proposers found the Release handle difficult to comprehend until the locking mechanism had been demonstrated. At the beginning of each trial, the locking mechanism would be activated, thus preventing the responder from acting until the proposer had made an offer. This forced the responder to wait their turn and not move until the proposer had made an offer, as is standard practice in all UGs.

It was explained to the participants that the more points they collected in their reward trays, the more prizes they would receive at the end and, conversely, any points that ended up in the black box would not count. Two responder habituation trials were then performed. The experimenter placed one point each at both proposer and responder's end of the playing surface and asked the responder to use their handle to tip the points into the reward trays (i.e. to the left). A second pair of points was then placed on the playing surface, again one at each end, but this time the responder was asked to tip them into the black box. The order of these two responder habituation trials was counterbalanced.

Once the responder's two choices had been demonstrated, attention shifted to the proposer. First, the experimenter locked the apparatus. The proposer was then told that their job was to "unlock the box", and that they could choose to do this in one of two ways. Four points were placed on the playing surface, two on the proposer's side and two on the responder's, one of which was in the contestable area and one in the non-contestable area (in other words, this was the set-up for the "selfish theft" condition show in Table 2). The proposer was then asked to pull the *Release* handle, and it was demonstrated to the responder that the box was now unlocked. The proposer's *Theft* handle was then demonstrated in the same manner, with four points arranged in the manner described above (selfish theft). Both proposer and responder were required to indicate which point had moved, where it had moved from (i.e. the

contestable area) and where it had moved to (i.e. the proposer's end of the apparatus. The contrast between the points in "the silver bit" (the contestable area) and "the green bit" (the non-contestable area) was emphasised.

Finally, the proposer was required to verbally explain the function of his or her two handles to the experimenter, one at a time (order counterbalanced across subjects). Proposers were first asked to explain what each did in their own words and then prompted to point to the one they should pull if they wanted to "leave things as they are," and the one they should pull if they wanted "to tip the box." If any important details were omitted, such as the difference between the contestable and non-contestable areas, questions were asked to prompt explanation. If the proposer failed to demonstrate understanding at any point, their habituation phase was repeated from the start. The responder was then asked to explain the different consequences of moving their handle to the left (reward) and right (punishment).

Test phase

Testing began immediately after the completion of the habituation phase. During each trial, the experimenter would place 4 points on the box's upper surface in the appropriate array and then prompt the proposer to "please choose one handle." Upon completion of the proposer's choice, the responder then would be asked "which way do you want to move the handle?" Each of the conditions in Table 2 was presented to all pairs once, in a pseudo-random, counterbalanced order.

Knowledge probes

Knowledge probes were conducted immediately after testing, in order to reconfirm participants' knowledge of the apparatus and procedure. The locations of the reward trays were reversed, so that each participant's reward tray was placed at their opponent's end of the apparatus. The purpose of this was to ensure that children were aware of the consequences of their actions upon their partner. They were told that their reward tray remained theirs, but that it had now been moved as part of "a little test." Participants were then individually asked twice (order counterbalanced) to operate their levers to move a ball at their partner's end (i.e. adjacent to the new location of their reward tray). Proposers were told that the experimenter would tip the responder's handle towards the reward trays and asked which handle they should pull in order to cause this to result in the ball falling in their reward tray. This new arrangement required proposers to inhibit pulling of the theft handle, which

had served to bring more rewards over to their side during the test phase. Responders were asked to move the ball at the other end once into their reward tray and once "to make it disappear" (i.e. to move it to the punishment tray).

3.3: Analysis

For proposers, the outcome variable was the frequency of Theft/Release choices, coded in a binary manner on each trial, based on which handle the proposer pulled. For responders the outcome variable was the frequency of Accept/Reject choices made in response to each offer, coded in a binary manner on each trial, based on whether the responder tipped the points into the reward trays or black box.

3.3.1: Analysis of Proposers

Firstly, a three-predictor stepwise logistic regression analysis was performed once for each of the seven conditions with theft (1) versus non-theft (0) as the outcome and age in months, gender and the relationship between participants (known=1, unknown=0) as predictors. This is described in results section 3.4.1, as *Proposer Analysis 1*. Any proposers who failed the knowledge probes were excluded from these regression analyses. This was done to ensure that only those who fully understood their choices were included in the data.

Participants were then grouped into the four age categories for further analysis, i.e. 4-5 years, 6-7 years, 8-9 years, 10+ years. Cochran's test was performed to check for overall differences in theft rate *between* conditions, but *within* each age category. For example, to compare different preferences between the seven conditions amongst pre-schoolers; then amongst 6-7 years olds, etc. If Cochran's test proved significant, McNemar's test was used for post-hoc analysis of conceptually interesting pairwise comparisons. First, zero-sum theft (0/4;0, i.e. where the proposer should always steal) was compared to the responder advantage control (0/0;4, i.e. where stealing achieves nothing). This served as a comprehension control, since theft in the former should be much higher, assuming proposers understood the effects of the theft handle. This stage of the analysis is described in section 3.4.1, *Proposer Analysis 2*.

Thirdly (Section 3.4.1: *Proposer Analysis 3*), binomial tests, using a comparison level of 50%, were conducted within each condition and within each age group in order to see whether a majority within each preferred the Theft or Release option. Finally, further McNemar tests

compared fair (1/1;2), selfish (2/1;1) and greedy theft (2/2;0), in order to discern the age at which proposers began to show different preferences towards offers which showed different levels of unfairness.

3.3.2: Analysis of Responders

A three-predictor stepwise logistic regression analysis was performed twice for each of the 7 conditions, once for responses to theft and once for responses to non-theft. In all 14 of these regressions, the same three predictors were used as in the proposer regressions (age, gender, relationship with opponent). This phase of the analysis is detailed in Section 3.4.2, *Responder Analysis 1*. Participants were then grouped into the same four age categories as the proposers (4-5 years, 6-7 years, 8-9 years, 10+ years). Within each condition and within each age category, two types of analyses were performed. Firstly, McNemar's test was used to determine whether the number of rejections differed significantly from null (*Responder Analysis 2*, Section 3.4.2), as in previous MUG studies with primates and young children (Jensen et al., 2007a; Kaiser et al., 2012; Wittig et al., 2013). Secondly, in conditions where rejection rates were significantly above zero, binomial analysis comparing against a baseline of 50% was used to test whether a majority of each sample preferred theft or rejection (*Responder Analysis 3*, Section 3.4.2).

3.4: Results

3.4.1: Proposers

Across the entire sample, the frequency of theft choices in each of the seven conditions was as follows: greedy theft (2/2;0) – 65%, fair theft (1/1;2) – 77%, pointless theft (2/0;2) – 31%, selfish theft (2/1;1) – 59%, modest theft (0/1;3) – 82%, responder advantage control (0/0;4) – 47%; zero-sum (0/4;0) – 83%. These figures show that theft levels were high. Indeed, theft was preferred by a majority in all conditions in which it could transfer extra points across to the proposer, and only in these conditions.

Proposer Analysis 1 – logistic regression

Thirty two proposers were excluded from the regression analysis after failing the knowledge probes. This left a total of 118 participants (mean age=9.19 years, SD=2.46, 67 girls, 51 boys) for logistic regression analysis of each of the seven conditions. The results of these regressions are presented in Table 3, which shows the predictors which significantly predicted Theft/Release in each condition.

Table 3: significant predictors of theft rates in Experiment 1.

condition	*predictors*	*β*	*Wald*	*p*
greedy theft (2/2;0)	age in months	-0.014	4.33	.037
selfish theft (2/1;1)	age in months	-0.014	4.76	.029
pointless theft (2/0;2)	age in months	-0.023	8.36	.004
zero-sum (0/4;0)	known>unknown	1.28	6.81	.009

As the table shows, increasing age in months predicted a decline in theft in three conditions: greedy theft (2/2;0), selfish theft (2/1;1) and pointless theft (2/0;2). Proposers were more likely to steal from people they knew than from strangers in the zero-sum condition. Aside from those presented in the table, there were no other effects of any predictors in any the conditions. This means that none of the predictors significantly influenced the theft rate in the fair theft, modest theft and responder advantage control conditions.

Proposer Analysis 2 – comparing theft rates within different age cohorts

Cochran's analysis was performed in order to check for differences in theft rates within the four age groups (4-5 years, 6-7 years, 8-9 years, 10+ years) but between conditions. No significant differences were observed amongst children aged 4-5 ($Q(6)=10.09$, $p=.121$), however, within all other age cohorts, the Q differential was extremely large (at 6-7 years, $Q(6)=43.86$, $p<.000001$; at 8-9 years, $Q(6)=31.69$, $p<.0001$; at 10+ years, $Q(6)=80.79$, $p<.000001$). This meant that proposers within the older three age cohorts were adjusting their theft rates according to condition.

Post-hoc analysis of key pairwise comparisons was performed via McNemar analysis. Zero-sum theft (0/4;0) was preferred to theft in the responder advantage control condition (0/0;4) by the three eldest age groups (6-7 years, $p<.0001$; 8-9 years, $p=.004$; 10+ years, $p=.001$). However, this distinction was not made amongst those aged 4-5 years ($p>.99$), amongst whom 84% preferred the theft handle even in the responder advantage control condition (0/0;4), again indicating that the youngest participants may not have understood when theft was and

was not possible. Having failed this control, children aged 4-5 years were excluded from the continuing analysis. All other age groups, having demonstrated no problem with this distinction, were included.

Proposer Analysis 3 – did a majority prefer theft in each condition?

Within the three remaining age cohorts, binomial analysis was performed, to see whether a significant majority favoured theft or non-theft. Figure 2 shows the percentage of thefts and the significance of the binomial analysis for each condition at each age group.

As the figure shows, at 6-7 years, children (n=31) demonstrated a significant preference for theft in all conditions in which it conferred a material advantage, but showed no preference for theft in the control conditions in which pulling the handle made no difference to their allocation. Children at this stage demonstrated preferences for selfish (2/1;1, proportion=.72, p=.005), zero sum (0/4;0, proportion=.91, p<.001), modest theft (0/1;3, proportion=.88, p<.001), greedy (2/2;0, proportion=.67, p=.032) and fair theft (1/1;2, proportion=.74, p=.002). Theft rates in the pointless theft (2/0;2) and responder advantage control (0/0;4) conditions revealed no significant preferences (proportion=.60, p=.22; proportion=.47, p=.76).

At 8-9 years (n=32), significant preferences for theft were again observed in many of the conditions in which it conferred an advantageous shift in resource allocation, with this phenomenon being observed in the greedy theft (2/2;0) proportion=.73, p=.006), fair theft (1/1;2, proportion=.73, p=.006), zero-sum theft (0/4;0, proportion=.75, p=.002) and modest theft (0/1;3, proportion=.73, p=.006) conditions. However, children by 8-9 years also demonstrated no actual preference on whether or not to engage in selfish theft (2/1;1; proportion=.56, p=.597). Also unlike the younger group, children at 8-9 years demonstrated a preference against pointless theft (2/0;2, proportion=.33, p=.038), although there remained no significant preference in the responder advantage control condition (0/0;4, proportion=.63, p=.15).

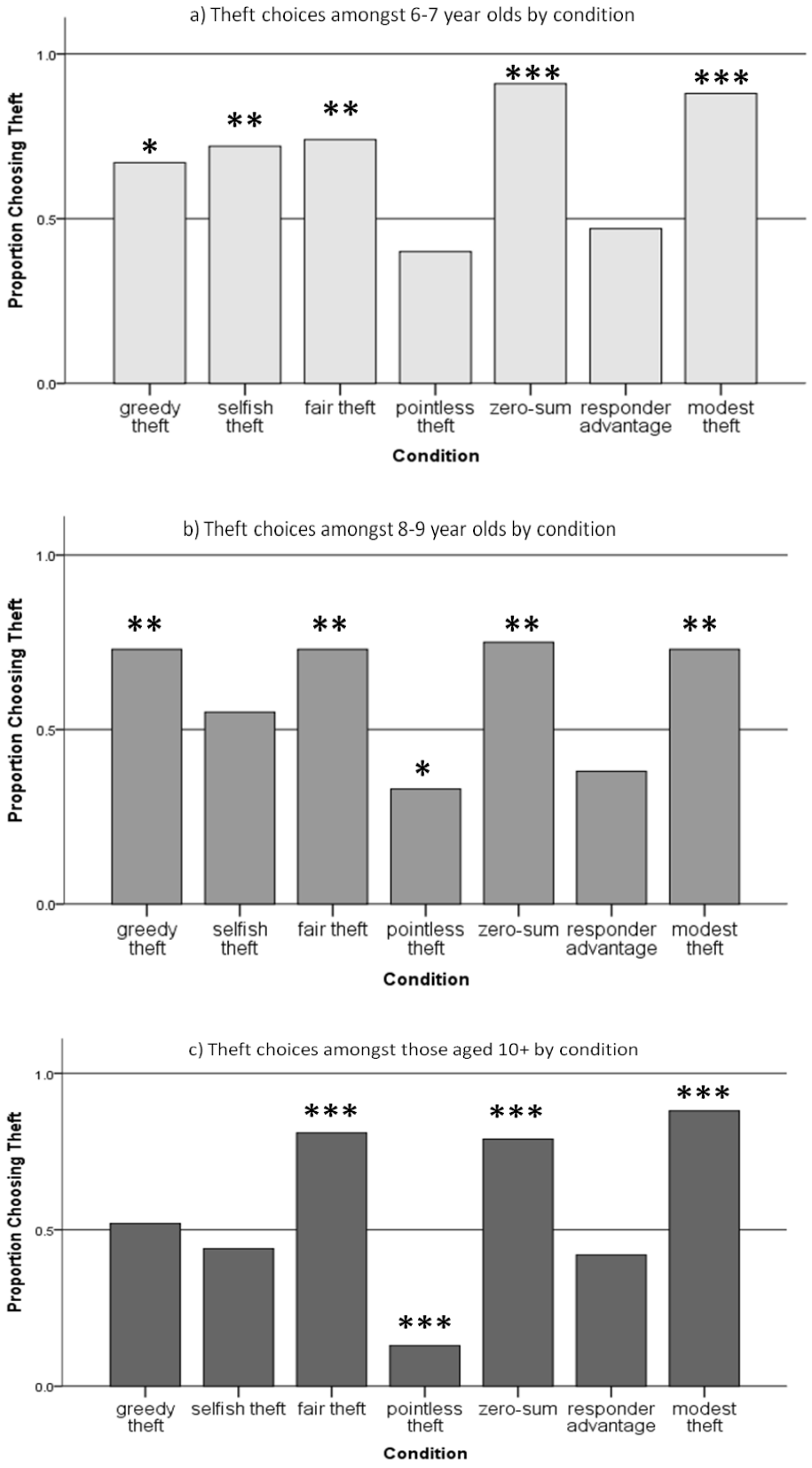

Figure 2: Proposer theft rates in Experiment 1.

Theft rates are shown by age cohort and condition, with binomial analysis indicating a preference for either theft or non-theft in each case: *p<.05, **p<.01, ***p<.001.

Only by 10+ years (n=45) were proposers no longer expressing a preference for obviously inequitable theft, with no significant majority in favour of greedy (2/2;0, proportion=.52, p=.885) or selfish theft (2/1;1, proportion=.44, p=.471). A significant preference for theft which conferred a material gain remained only in the cases of fair theft (1/1;2, proportion=.81, p<.001), zero-sum theft (0/4;0, proportion=.79, p<.001) and modest theft (0/1,3, proportion=.88, p<.001). As with children at 8-9 years, children aged 10+ preferred not to pull the theft handle in the pointless theft condition (2/0;2, proportion=.13, p<.001), but had no preference in the responder advantage control condition (0/0;4, proportion=.42, p=.233).

Finally, pairwise McNemar analyses were used to determine whether proposers could adjust their rates of theft according to how unfair the theft was in different conditions. Comparisons were made between offers which left responders with a fair (50%) share versus those in which they were left with nothing (fair theft, 1/1;2 vs. greedy theft, i.e. 2/2;0), offers which left responders with a small (25%) share versus those in which they were left with nothing (selfish theft, 2/1;1 vs. greedy theft, 2/2;0), and offers which left responders with a small share versus those in which they were left with nothing (selfish theft, 2/1;1 vs. greedy theft, i.e. 2/2;0). This was done in order to see if proposers appreciated that a greater likelihood of rejection was probable when they left the responder with small (25%) or, in particular, zero offers.

Proposers largely failed to make this distinction until they reached late childhood. Only by 10+ years did proposers prefer fair theft to greedy theft (1/1;2 over 2/2;0, p=.003), and fair theft to selfish theft (1/1;2 over 2/1;1, p=.001), with all other comparisons at all age groups showing an inability or unwillingness to refrain from the comparatively selfish offer relative to the comparatively generous one (all p-values >.25).

3.4.2: Responders

All responders correctly answered the knowledge probes, thus, unlike the proposer dataset, there were no exclusions from the final analysis (n=150). Since spiteful responding was the chief focus of this experiment, responders were included in the analysis even when teamed with proposers who failed the knowledge probes and consistently pulled theft handle. Overall rejection rates of the offer caused by proposer *theft* were as follows: greedy theft (2/2;0) – 70%, fair theft (1/1;2) – 17%, pointless theft (2/0;2) – 17%, selfish theft (2/1;1) – 42%, modest theft (0/1;3) – 14%, responder advantage control (0/0;4) – 16%; zero-sum (0/4;0) – 69%. This

means that only offers of zero were rejected by a majority of subjects, and that any offer of 50% or more was accepted by a large majority.

Responder Analysis 1 – logistic regression

Stepwise logistic regressions were used to examine whether age in months, gender and relationship predicted rejection of offers *in response to theft only*, i.e. the figures quoted in the previous paragraph. The results of these regression analyses are presented in Table 4. Increasing age in months predicted higher rejection rates towards the lower offer in the greedy theft (2/2;0), selfish theft (2/1;1) and zero-sum (0/4;0) conditions. These were the only significant findings, meaning that none of the demographic variables predicted rejection frequencies in any of the other conditions.

Table 4: significant predictors of offer rejection in Experiment 1.

condition	*predictors*	*β*	*Wald*	*p*
greedy theft (2/2;0)	age in months	-0.017	3.82	.05
selfish theft (2/1;1)	age in months	-0.021	6.82	.009
zero-sum (0/4;0)	age in months	-0.2	6.54	.011

In the second set of stepwise logistic regressions looking at responses to *non-theft* only, none of the predictors significantly affected rejection rates in any of the seven conditions. In other words these offers, all of which consisted of 50-100%, were simply accepted at a uniform rate by all subjects.

Responder Analysis 2 – did rejection rates differ from null?

Participants were then grouped into the four age categories (4-5 years, 6-7 years, 8-9 years, 10+ years). Within each condition and within each age category, McNemar's test was used to determine whether the number of rejections differed significantly from null. Data showing rejection rates towards theft in each condition at each age group is shown in Table 5.

Table 5: responder violations of the SEM in Experiment 1.
*McNemar analyses show whether rejection rates differed from zero within each age cohort: *p<.05, **p<.01, ***p<.001.*

		condition and notation						
age group	preferences	greedy theft 2/2;0	fair theft 1/1;2	pointless theft 2/0;2	selfish theft 2/1;1	modest theft 0/1;3	responder advantage 0/0;4	zero-sum 0/4;0
4-5 years	% favouring rejection	32%	14%	25%	19%	15%	20%	48%
	different from null?	yes*	no	no	no	no	no	yes***
6-7 years	% favouring rejection	79%	27%	19%	31%	16%	13%	65%
	different from null?	yes***	yes**	no	yes**	no	no	yes***
8-9 years	% favouring rejection	91%	15%	0%	55%	7%	11%	81%
	different from null?	yes***	no	no	yes***	no	no	yes***
10+ years	% favouring rejection	76%	12%	17%	67%	16%	20%	81%
	different from null?	yes***	no	no	yes***	yes*	no	yes***

The table shows several deviations from the SEM in responder decisions, although on the whole these represented a relatively rare occurrence. The most important column is the selfish theft condition, in which the Theft handle made an offer (25%) which was greater than zero but clearly disadvantageous when compared to the proposer. Here, 4-5 year olds (pre-schoolers) failed to deviate from the SEM, with a theft rate not significantly different from null (p=.125). All of the older age groups, however, did reject 25% at a rate significantly above null in this condition. There were few other deviations from the SEM on display, but this is to be expected, since in the other conditions in which theft left the responder with a proportion of the allocation, this proportion was always 50% or over. Rejections of zero offers in the greedy theft (2/2;0) and zero-sum (0/4;0) conditions were, unsurprisingly, consistently higher than null.

Responder Analysis 3 – did subjects favour acceptance or rejection of theft?

The previous section used the SEM as a baseline against which to test children's actions. For example, Table 5 shows that 31% of 6-7 year olds in the selfish theft condition (2/1;1) preferred to reject the theft offer (1 point), and that this percentage is significantly higher than zero, thus violating the SEM.

This is interesting, but 31% is still a modest figure – clearly only a minority of children at this age were willing to violate the SEM. This section therefore aims to go further by examining whether subjects formed clear majorities in favour of rejection or acceptance in each case. Binomial analyses against a baseline of 50% were applied to the same data on responder

rejection rates at each age, to test whether significant majorities of the sample favoured acceptance or rejection. So, to return to the previous example, if 31% of 6-7 year olds favoured rejection in the selfish theft condition, this means that 69% preferred to accept the offer of 1 point. Does this 69% constitute a significant majority of subjects, i.e. does the overall sample prefer to accept theft in this case? In this case, the answer was no (p=.076), thus demonstrating that children at 6-7 years had no significant preference as to whether or not to accept offers of 25%, even when it was the product of theft. The results of this response and all of the other analyses are presented in Table 6.

Table 6: responder rejection preferences in Experiment 1.
*Binomial analyses show whether responders prefer to accept or reject theft by condition and age cohort: *p<.05, **p<.01, ***p<.001.*

		condition and notation						
age group	preferences	greedy theft 2/2;0	fair theft 1/1;2	pointless theft 2/0;2	selfish theft 2/1;1	modest theft 0/1;3	responder advantage 0/0;4	zero-sum 0/4;0
4-5 years	% favouring rejection	32%	14%	25%	19%	15%	20%	48%
	different from 50%?	no	accept***	no	accept**	accept***	accept*	no
6-7 years	% favouring rejection	79%	27%	19%	31%	16%	13%	65%
	different from 50%?	reject**	accept*	accept**	no	accept***	accept***	no
8-9 years	% favouring rejection	91%	15%	0%	55%	7%	11%	81%
	different from 50%?	reject***	accept***	accept***	no	accept***	accept***	reject**
10+ years	% favouring rejection	76%	12%	17%	67%	16%	20%	81%
	different from 50%?	reject*	accept***	accept*	no	accept***	accept*	reject***

At 4-5 years children did not prefer to reject offers in any condition, even those in which they were left with zero. Instead, in 4 out of the 5 conditions in which they were left with at least part of the endowment (including 25% in selfish theft), a significant majority chose to accept (all p-values <.04). They were also indifferent as to whether to accept offers of zero in the greedy theft and zero-sum conditions, despite the fact that the proposer was left with 100% in both of these conditions. In other words, children of a pre-school age adhered perfectly to the SEM in this experiment.

By 6-7 years, a significant majority preferred to reject zero in cases of greedy theft (2/2;0, p=.002), and marginally so in cases of zero-sum theft (0/4;0, p=.081). The only majorities in favour of acceptance were in cases where they were faced with offers of at least 50% (fair theft 1/1;2, p=.014; pointless theft 2/0;2, p=.007; responder advantage control 0/0;4, p<.001; modest theft 0/1;3, p<.001). Offers of 25% in the selfish theft (2/1;1) condition had shifted from a significant majority in favour at 4-5 years, to a marginal trend against (p=.076).

By 8-9 years, responders retained their majority preferences for accepting offers in conditions in which offered at least 50% (all p-values <=.001), and they also now strongly preferred to reject both offers of zero (greedy theft 2/2;0, p<.001; zero-sum 0/4;0, p=.002). This pattern was almost identical at 10+ years.

3.5: Discussion

Four hypotheses were tested in this Experiment. **Hypothesis 1**, that the SEM would be violated at all ages, was not supported, as while the eldest three age cohorts all demonstrated negative other-regarding preferences, pre-school children did not. **Hypothesis 2** was supported: costly second-party punishment did increase in conjunction with age. **Hypothesis 3** was partially supported. Children of all ages overwhelmingly preferred to accept fair and advantageous offers, and there was no linear fluctuation in this preference according to age. However, contrary to H3, cost-free second-party punishment was observed to increase with age, as pre-school children demonstrated indifference rather than punitive preferences in response to offers of zero. **Hypothesis 4** was also partially supported. Proposers did make fewer unfair offers with increasing age. However, proposer theft rates were on the whole fairly high, and there was no obvious switch to an active preference for fair distributions in those aged 8-9 years and above in cases where the alternative was *advantageous* inequity.

3.5.1: Responder Discussion

Generally speaking, negative other-regarding preferences became more pronounced with age, and more skilfully targeted towards appropriate contexts. In common with Wittig et al (2013) and Bereby-Meyer and Fiks (2012), but in contrast to Murnighan and Saxon (1998) and Takegishi et al. (2010), this experiment found pre-school responders behaving very close to the SEM in an MUG, with little evidence of negative other-regard or hyper-competitive behaviour. This is clear from their consistent preferences for accepting all offers of greater than zero, and the fact that all offers of greater than zero, including the offer of 25% in the selfish theft (2/1;1) condition, were not rejected by a proportion significantly greater than null. These preferences are virtually identical to those of our closest primate relatives in the MUG (Jensen et al., 2007a; Kaiser et al., 2012; Proctor et al., 2013), and indeed to the predictions of the SEM.

Perhaps an even more vivid example of a lack of negative other-regard is provided by the fact that pre-school children exhibited no firm preference for or against rejecting offers of *zero* (i.e. greedy theft, 2/0;2, and zero-sum theft, 0/4;0). In other words, even in cases where they had been clearly wronged and punishment was cost-free, responder choices appeared to reflect indifference rather than punitive anger. This finding was contrary to predictions, and is surprising in light of several previous studies. Children aged around 4-5 have previously been observed paying a spiteful cost to remove disadvantageous inequity (Blake & McAuliffe, 2011; Bügelmayer & Spiess, 2014; McAuliffe et al., 2013; McAuliffe et al., 2014; Sheskin, Bloom, et al., 2014), so their lack of consistent punishment in the present experiment even when no cost was required was unexpected. When the cost of punishment is not a factor, pre-school children have even been previously observed enacting third party punishment on others' behalf (Kenward & Östh, 2012; Riedl et al., 2015), so to see many of them failing to do so on their own behalf here was unexpected

In line with expectations costly punishment did increase monotonically with age in the selfish theft condition. On average, older children were more likely to display negative other-regarding preferences in a way which violated the SEM. It is noteworthy, however, that *non*-costly punishment also increased with age, even though children in the younger age brackets of the sample were of an age at which an ability to understand the fairness norm has been previously demonstrated (Smith et al., 2013; Steinbeis et al., 2012). These results suggest that a preference for punishments of all types increases throughout childhood to early adolescence. It also hints, like other recent work (e.g. Kogut, 2012; Smith et al., 2013) that adherence to the fairness norm is not of great importance to many children at pre-school age. This aspect of the results in still surprising, however, in light of recent findings that younger children have higher levels of negative other-regarding concerns such as envy and schadenfreude (Fehr et al., 2013; Steinbeis & Singer, 2013). Even if these were suppressed in the selfish theft condition (2/1;1) due to a desire to self-maximise and accept the point on offer, they would have been expected to motivate cost-free punishment in the greedy (2/0;2) and zero-sum (0/4;0) theft conditions.

It is very important to note at this point that, unlike the youngest proposers, there is little doubt that responders at 4-5 years understood the apparatus. Mechanistically, their task was far simpler than that of proposers, something reflected in the fact that none of the 150 tested failed the responder knowledge probes. Also important to note is that at no point did the number rejecting the responder advantage control (0/0;4) of 100% of the allocation differ

significantly from null, meaning that virtually all children of all ages preferred to accept 100% all of the time, and that offers of 100% were unique in this respect. This is further strong evidence that responders both understood the scenario and were motivated by it, as, in a binary choice paradigm, virtually any random behaviour at all would have nullified this result in such a large dataset. Furthermore, in line with predictions, children of all ages consistently expressed majority preferences for accepting all fair or advantageous offers, thus strongly suggesting that they were motivated to maximise their income when doing so did not put them at a relative disadvantage. It also shows that children at all ages considered fair and advantageous offers to be acceptable, as predicted.

Between 4-5 years and 6-7 years, a radical change in negative other-regard appeared to occur, as shown by the sharp change in the pattern of punitive responses. Although a slightly more targeted attitude to rejection continued to be refined with age (for example, with rejections differing from null only in cases of offers of 0-25% at 8-9 years, rather than offers of 50% at 6-7 years), the most notable change in behaviour occurred in the shift between the youngest two age groups. By 6-7 years, clear violations of the SEM were now present in the form of significant costly punishment (e.g. fair theft, selfish theft), and strong majorities favoured cost-free punishment, in contrast to the indifference seen in 4-5 year olds. This pattern remains broadly similar for the rest of childhood, albeit with the odd idiosyncrasy fading with age, such as the disappearance of a minority in favour of rejecting equal offers in the fair theft condition (1/1;2). Indeed, the punitive behaviour of the elder three age groups is, on the whole, rather similar. This too is an important finding, since it perhaps goes some way to explaining why previous ultimatum games focusing only on children from mid-childhood through to adolescence have not always found clear linear increases in rejections at such ages (e.g. Sutter, 2007). It may be that the most important growth in punitive sentiment occurs prior to this age. A similar conclusion was reached in another recent study of children's negative other-regarding preferences in the MUG (Bereby-Meyer & Fiks, 2013).

While it is true that at no age did a significant majority ever favour rejecting offers of 25%, this finding is not necessarily indicative of weak other-regarding preferences, although it does suggest that participants were not so strongly fixated on simple disadvantageous inequity aversion that they would automatically pay to punish anything which deviated from it. The adults in Falk, Fehr and Fischbacher's original (2003) MUG study rejected offers of 20% in the equivalent condition at a rate similarly close to 50% and, indeed, a review of the UG literature on the whole shows that adult samples also tend to be split as to whether or not to accept offers of approximately this size (Camerer, 2003). The responders in Experiment 1, at least in

the case of those aged 6-7 years and upwards, simply appear to have been divided as to whether 25% is an acceptable offer in the UG, as are many adults. It is true, however, that in the present experiment the responders in middle to late childhood (those in the elder two age cohorts) appear to have been slightly more conservative in their rejection rates of offers of 25% than were children and teenagers in the equivalent condition in some other developmental ultimatum games (Güroğlu et al., 2009; Sutter, 2007).

Little evidence of advantageous inequity aversion was apparent in the responders' behaviour, despite this being a behaviour which has been demonstrated before in children aged as young as 4-8 years (Blake & McAuliffe, 2011; McAuliffe et al., 2013; McAuliffe et al., 2014). The offer of 100%, as discussed, was never rejected, even by a minority, at a rate significantly different to null. The 75% offer (i.e. in modest theft, 0/1;3) was rejected by a small but significant minority at 10+ years, although it is not wholly clear if this was due to advantageous inequity aversion, to a punitive reaction from a minority of subjects who were particularly averse to theft, or from a mixture of the two.

The fact that fair theft (1/1;2) was not rejected by the same age group does perhaps hint at some degree of advantageous inequity aversion here, as if it was merely anger towards theft causing rejection then fair theft ought to have been rejected at a rate higher than null too, given that doing so was less costly (costing two tokens as opposed to three). However, as Shaw and Olson (2012) correctly point out, rejecting advantageous inequity, as opposed to disadvantageous inequity, is often more costly in absolute as well as relative terms in economic experiments of this type (a critique which applies to the present study), so comparing the two directly in responder behaviour is problematic. For example, a rejection of advantageous inequity in the present study would have cost the responder 3-4 tokens, but a rejection of disadvantageous inequity would only have cost the responder one token. Thus, it is perhaps unsurprising that responders were less willing to choose the former, even if some of them would have theoretically preferred an even split if given the choice. Better evidence of the relative absence of an aversion to advantageous inequity in the present experiment is provided by the proposer data.

3.5.2: Proposer Discussion

In line with expectations, proposers were more likely to make fair offers the older they got in key conditions. These were selfish theft (2/1;1) and greedy theft (2/0;2), in which choosing

theft became less likely the older the proposer was. Therefore, in the language of ultimatum games, offer size increased in tandem with age. These basic trends were largely in line with expectations, and yet the proposer results remain somewhat surprising in the context of previous experimental work on children's fairness concerns. This is because many proposers were strikingly ungenerous, with many making grossly unfair offers both at and well above the age of 8-9 years, at which previous studies have found a firm behavioural preference for fairness (Blake & McAuliffe, 2011; Fehr et al., 2008; House et al., 2013; McAuliffe et al., 2013; McAuliffe et al., 2014; Smith et al., 2013). Indeed, although the "theft" aspect of our design was primarily aimed at prompting rejection from responders, it appears to have had a more pronounced effect in prompting selfish behaviour from proposers.

There are two particular aspects in which the proposer results were less generous than predicted. Firstly, as discussed, children have frequently been observed to exhibit preferences for fair distributions over advantageous inequity by 8-years of age in many influential existing studies (Blake & McAuliffe, 2011; Blake, McAuliffe, & Warneken, 2014; Fehr et al., 2008; House et al., 2013; Kogut, 2012; McAuliffe et al., 2013; McAuliffe et al., 2014; Sheskin, Bloom, et al., 2014; Smith et al., 2013). In the present experiment, however, evidence of an aversion to advantageous inequity was absent throughout the sample. By 10+ years children had no active preference in favour of selfish or greedy acts of theft, and did at least succeed in inhibiting both choices as compared to acts of fair theft (equal offers). However, this is not the same thing as having an active preference *in favour* of fair offers over advantageous inequity (i.e. a preference for not stealing in the fair theft and greedy theft conditions), something which previous research suggests should have been present by 8-years (e.g. Blake & McAuliffe, 2011), but which was in fact absent throughout the entire sample, which included many participants well above this age. Indeed, at 8-9 years, the age at which the key shift towards fair outcomes has previously been observed, children actively preferred greedy theft, the single most egregiously unfair act in the entire experiment. Below this age, all thefts which conferred any kind of material advantage were preferentially chosen, a behaviour which is again very reminiscent of UG experiments with higher primates (Jensen et al., 2007a; Kaiser et al., 2012).

A second unusual aspect of these findings is that our proposers' behaviour was not just selfish, but also strategically very naïve. Evidence of sophisticated strategic offer-making behaviour in variants of the UG has previously been observed to be absent in pre-school children (Wittig et al., 2013), but present in those from middle-childhood upwards (Harbaugh et al., 2003; Kogut, 2012; Steinbeis et al., 2012). Thus, even if bereft of egalitarian sentiment, by 8+ years children

should have been able to draw the conclusion that offering a responder 25% was, at best, risky, and offering them zero was extremely unlikely to meet with acceptance. Indeed, as responders, younger children than this (those from 6-7 upwards) routinely and preferentially rejected offers of zero, thus demonstrating that an understanding that this behaviour was unfair and deserving of punishment was clearly present in children aged much younger than many of those making such offers (e.g. 52% of the sample, even in those aged 10+).

Nevertheless, it should be noted that by 10+ years, although not before, children were only expressing significant majority *preferences* for theft in the three conditions in which it was arguably justifiable according to conventional notions of fairness, i.e. fair theft (in which they were evening a disadvantageous distribution), zero-sum theft (in which no sharing was possible) and modest theft (in which they started with nothing and took only 25%). Offers did increase in size in conjunction with age, a result which accords with previous developmental work in both ultimatum games specifically (Harbaugh et al., 2003; Kogut, 2012; Steinbeis et al., 2012). One issue with the proposer data which must be highlighted, however, is that the knowledge probes and Cochran tests showed that the youngest proposers could not reliably demonstrate full comprehension of their role. The apparatus may have been too complex for them, or they may simply have been unable to resist pulling the potentially more rewarding theft handle. Regardless of the reason, conclusions on proposer offers can only reliably be drawn on those aged 6 and upwards.

3.5.3: Conclusion

Experiment 1 represents a largely successful application of a novel version of the MUG to a human sample for the first time. It has unearthed a large array of original data with important implications for the development of children's negative ORPs. Although the general observation that these were increasingly common towards the elder end of the sample was to be expected, other findings clearly require further investigation. From the point of view of the emergence of hyper-competitiveness, three findings in particular stand out. First is the lack of evidence of hyper-competitive behaviour as represented by the low levels of punitive behaviour in pre-schoolers. Second is the large increase in negative ORPs between 4-5 years and 6-7 years. Third is the increase in preferences for both costly and cost-free punishment with age. Also of interest is the fact that many proposers exhibited extremely unfair and competitive behaviour at well past the point in development where previous work had suggested they should cease to do so. It is very likely that the "theft" framing of the

experiment played a strong role in encouraging the latter in particular. Given the importance of the key findings and the general lack of reliability in developmental UG studies, a replication and revision of the MUGWT will also be employed in Experiment 2. In particular, an attempt will be made to expose responders to more scenarios in which they are presented with borderline acceptable offers, in order to replicate and extend key findings regarding the emergence of costly second-party punishment.

Chapter 4: The mini-ultimatum game with theft with revised conditions

4.1: Introduction

The first application of the MUGWT to human subjects in Experiment 1 can be considered a qualified success. However, the key focus of this dissertation is negative other-regarding preferences, and particularly instances in which they cause the actor to pay a direct cost. Given this focus, one potential drawback to Experiment 1 was that it was comprised of seven conditions, but only one of these (i.e. selfish theft, 2/1;1) involved a scenario in which a responder was presented with a theft offer which was both unfair, in that it involved disadvantageous inequity, and yet also greater than zero, and thus necessitated paying a cost to reject. Although Experiment 1's findings relating to the costly punishment of disadvantageous inequity were important, Experiment 2 will seek to extend them considerably by increasing the number of instances in which offers of 25% are presented to responders, in order to examine this most important scenario in more depth. The greater range of scenarios in which small offers can be made to the responder will also allow for inferences to be drawn as to whether children can differentiate between outcome and intention (Falk et al., 2003) when rejecting unfair offers, or instead simply make their choices based on inequity aversion (Fehr & Schmidt, 1999; Loewenstein et al., 1989). Finally, as the introduction to the previous chapter made clear, the reliability of findings between different developmental UG studies is often low, so this chapter will also aim to replicate the key findings of Chapter 3.

4.1.1: Aims and Objectives

The aims and objectives of Chapter 4 are as follows:

1) To examine once again pre-schoolers' willingness to violate the SEM as responders in the UG, this time across a wider range of contexts
2) To gain insight into whether responders of all ages, but particularly towards the higher end of the age spectrum tested, differentiate between the outcome of proposers' offers, and the intention that underlies them.

3) To attempt to replicate the key responder findings from Experiment 1, specifically the lack of ORPs in pre-school children and the monotonic increase in costly and non-costly rejections of disadvantageous inequity with age.
4) To replicate the key proposer findings from Experiment 1, specifically the increase in fair offers at the expense of advantageously unfair offers with age, and the relatively high rates of overall theft.

4.1.2: Revised Aspects of the MUGWT Methodology in Experiment 2

In order to achieve its aims, Experiment 2 involved a revised raft of conditions, this time based on Falk et al's original (2003) MUG, as opposed to Kaiser et al.'s (2012) MUGWT. In the latter, and therefore also in Experiment 1, all of the experimental conditions (conditions 1-4 in Table 2, see Chapter, 3, Section 3.2.3, p.55) involved choices between an equal 50/50 split, and an unequal alternative of some description. The 50/50 split was therefore a constant feature of the experimental conditions. This was done in order to ease comprehension of the comparison between the "fair" and the alternative choices as much as possible. In the original MUG (Falk et al., 2003), however, the recurring baseline offer in the experimental conditions is one which is unequally split 80/20 in the proposer's favour. This has been the standard arrangement in MUGs since (Bereby-Meyer & Fiks, 2013; Castelli, Massaro, Bicchieri, Chavez, & Marchetti, 2014; Gummerum & Chu, 2014; Güroğlu et al., 2009; Sutter, 2007; Takagishi et al., 2010; Takagishi et al., 2014; Wittig et al., 2013), and it is therefore the one which Experiment 2 will adopt. Doing so will expose responders to four conditions in which they can potentially be made an offer which is both unequally balanced against them and yet also costly to reject. In any of these conditions, making the choice to reject violates the SEM, and is thus an example of a hyper-competitive, negative other-regarding preference.

4.1.3: Proposers' Outcomes and Proposers' Intentions

Experiment 2 will also provide some insight as to whether pre-adolescent children can differentiate between the fairness of an outcome and the fairness of the underlying intention behind it. This distinction was a central focus of the original MUG (Falk et al., 2003). Children will this time be presented with multiple offers of 25%, some of which are fairer in relative terms than the alternative, and others which have the same or a lesser degree of relative fairness. For example, when the alternative is zero (as it is in the greedy theft 2 condition), 25% is the fairest offer the proposer can make. When the alternative is 50%, however (as it is in the

selfish theft condition), 25% is an unfair offer. The outcome is the same in both conditions. The intention as to whether to treat the responder fairly is not.

Whether or not children are willing or capable of targeting their responses in each case will be interesting to see. It is something towards which I retain an open mind, since previous developmental MUG studies have reached differing conclusions as to whether or not pre-adolescent children's response preferences distinguish between these two constructs. To date, two studies have suggested that they can (Bereby-Meyer & Fiks, 2013; Sutter, 2007), and three that they cannot (Castelli et al., 2014; Gummerum & Chu, 2014; Güroğlu et al., 2009). One thing which all of these studies share, however, is a focus on middle to late childhood (typically around 8-9 years upwards) as the age at which the outcome/intention distinction might realistically appear. Success in understanding this relatively subtle distinction has not been observed in children younger than this. It should be stressed, however, that evidence from both the developmental UG (Kogut, 2012; Steinbeis et al., 2012) and other experiments (Smith et al., 2013) does show that younger children *understand* the distinction between fair and unfair offers just as clearly as older children. Nevertheless, it seems that they may apparently lack the motivation to respond spitefully to the latter to the same extent as elder children and adults. Their ORPs appear to be less strong in this respect.

Conflicting results, such as those on outcome/intention understanding outlined in the previous paragraph, provide a good example of how the reliability of UG and MUG findings with developmental samples is curiously low from one study to the next. For this reason, a secondary aim of Experiment 2 is to seek to replicate and confirm the key findings of Experiment 1. These include pre-school children's reluctance to violate the SEM, the increase in negative other-regarding preferences with age, and a lack of fairness preferences amongst proposers, as evidenced by a high frequency of selfish offers.

4.1.4: Hypotheses

On the basis of the aims and areas of interest discussed, this chapter will test the following hypotheses:

- **(H1)** In replication of E1, those aged 6-7 years and over will violate the SEM by displaying costly, second-party punishment. Pre-school children, however, will not.

- **(H2)** In replication of E1, the frequency with which children prefer costly second-party punishment of disadvantageous inequity will increase monotonically with age. Older children will be more willing to violate the SEM.
- **(H3)** In replication of E1, older children will also be more willing to enact cost-free punishment. The pre-school cohort, unlike the older children, will be indifferent to cost-free punishment.
- **(H4)** In replication of E1, responders across the age spectrum will be consistent in their preferences towards other kinds of offers. Acceptance rates of fair and advantageous offers are expected to be consistently high.
- **(H5)** Towards the upper end of the age spectrum, proposers will start to show an understanding of the distinction between outcome and intention.
- **(H6)** Proposer thefts will become less frequent with age in conditions in which they cause an offer which is unfair to the responder.
- **(H7)** In replication of E1, however, proposer theft rates will remain high, and at no stage of ontogeny will proposers actually *prefer* generous offers.

4.2: Methods

Experiment 2 largely represented a replication of Experiment 1, therefore many aspects of the methodology were identical to those outlined in the previous chapter, including the apparatus and procedure. Different aspect of the sample and design are outlined below.

4.2.1: Participants

Recruitment and sampling methods were the same as in Experiment 1. The sample consisted of 330 children aged 4-13 years, who were divided for analytical purposes into the same age categories as in Experiment 1. Of the 165 proposers (mean age=8.75 years, SD=2.42), 83 were female and 82 male. Of the 165 responders, 74 were female and 91 male (mean age=8.5 years, SD=2.7). During analysis the sample was broken down into four categorical age cohorts. The number of proposers within each of these cohorts was as follows: 4-5 years, n=23; 6-7 years, n=40; 8-9 years, n=46; 10+ years, n=56).

4.2.2: Design

As in Experiment 1, seven conditions were included in the experiment, and these once again included the three novel control conditions introduced to Experiment 1 (conditions 5-7, see Table 7). Conditions 1-4 were this time derived from those used in the original MUG (Falk et al., 2003). A full explanation of all seven conditions is provided by Table 7. Note how the possibility of a 3-1 rather than a 2-2 division is now the constant in all of these first four conditions. As a consequence, the "fair theft" condition, in which taking caused a 2-2 split, has now been replaced with a "justifiable theft" condition, in which an even split is impossible, leaving the proposer forced to make an unequal division to the advantage of one player or the other. "Greedy theft 2" and "pointless theft 2" are, as their names suggest, still conceptually similar to their equivalent conditions in Experiment 1, but feature 3-1 as the default division instead of 2-2. The notation used to explain these conditions is the same as in the previous chapter (see Chapter 3, Section 3.2.3 for a full explanation). Consistent with Experiment 1, the sample was divided, between-subjects, into an equal number of proposer/responder pairs, and all pairs experienced all seven conditions in a counterbalanced, pseudo-random order.

Table 7: the seven conditions in Experiment 2.

condition	*notation*	*initial distribution*	*distribution after theft*
1. greedy theft 2	3/1;0	3 - 1	4 - 0
2. justifiable theft	1/2;1	1 - 3	3 - 1
3. pointless theft 2	3/0;1	3 - 1	3 - 1
4. selfish theft	2/1;1	2 - 2	3 - 1
5. modest theft	0/1;3	0 - 4	1 - 3
6. responder advantage	0/0;4	0 - 4	0 - 4
7. zero-sum	0/4;0	0 - 4	4 - 0

4.2.3: Procedure

The procedure was virtually identical to that used in the previous chapter (see Chapter 3, Section 3.2.4 for full details), with the obvious exception of the fact that it used the revised conditions as detailed in Table 7. As in Experiment 1, subjects went through a familiarisation phase, a test phase, and a knowledge probe phase, all of which they experienced in face-to-face pairs as either proposer or responder. During the familiarisation phase the responder's role was explained first, via two counterbalanced trials demonstrating how to use the apparatus to accept and reject outcomes. The proposer's role was then explained by two counterbalanced trials demonstrating how to make offers via either theft or non-theft. Both

subjects were required to explain these roles verbally back to the experimenter before the experimental trials began. All subjects were required to attempt to demonstrate their continued understanding of the apparatus post-experiment via the knowledge probe trials. At the end, participants were rewarded with prizes purchased via the points they accumulated on the experimental trials, in the usual manner.

4.2.4: Analysis

All of the analyses reported in Chapter 3 were repeated in Chapter 4. The six subsections in the following results section are therefore the same as in Experiment 1, and are explained in Sections 3.3.1 and 3.3.2 of the previous chapter.

In addition to these same analyses, responder data in Experiment 2 for responses to *non*-theft were included in two conditions, namely greedy theft (3/1;0) and pointless theft 2 (3/0;1). This was done because in these two conditions non-theft now presented the responder with the combination of both disadvantageous inequity *and* an offer of greater than zero. Consequently, the data in these conditions were included at all sub-stages of the responder analysis section.

4.3: Results

4.3.1: Proposers

Amongst the entire sample, the frequency of theft choices in each of the seven conditions was as follows: greedy theft (2/2;0) – 50%, justifiable theft (1/2;1) – 81%, pointless theft (3/0;1) – 22%, selfish theft (2/1;1) – 56%, modest theft (0/1;3) – 78%, responder advantage control (0/0;4) – 42%; zero-sum (0/4;0) – 82% . Theft levels were therefore again fairly high, if not quite as high as in Experiment 1 in the altered greedy theft condition (3/0;1). The theft option was preferred by at least half of subjects in all conditions in which it could transfer extra points across to the proposer, and only in these conditions. This finding is very similar to its equivalent in Experiment 1.

Proposer Analysis 1 – logistic regression

Twenty one proposers were excluded from the analysis after failing the knowledge probes. This left a total of 144 participants (mean age=9.09 years, SD=2.35, 71 girls, 73 boys) for analysis. Logistic regression analysis, looking for the influence of demographic factors on theft rates, found fewer linear age effects than in Experiment 1. There was only a marginal decline in theft in the responder advantage control condition (0/0;4, β=1.076, Wald=3.74, p=.053). There were no effects of age in any of the other conditions, and no effects of relationship or gender in any of the conditions.

Proposer Analysis 2 – comparing theft rates within different age cohorts

Cochran's analysis found significant differences in theft rates between the seven conditions at all age groups. Unlike in Experiment 1, a significant Q differential was observed amongst children aged 4-5 years (n=16, Q(6)=15.506, p=.017), as well as amongst the older groups aged 6-7 years (n=32, Q(6)=37.398, p<.001), 8-9 years (n=40, Q(6)=55.131, p<.001) and 10+ years (n=56, Q(6)=83.516, p<.001). This means that there were significant differences in the theft rate across the seven conditions within all four age cohorts.

Again, as post-hoc analysis, the most important pairwise McNemar comparisons were deemed to be between the responder advantage control condition (0/0;4), in which theft conferred no difference in allocation, and the zero-sum theft conditions (0/4;0), in which a proposer should be expected to steal for both selfish and tactical reasons. Analysis showed that, once again, children at 4-5 years failed to choose theft preferentially in the zero-sum condition (p=.453), suggesting an inability to understand when theft was and was not possible. Amongst all of the older age categories, all of these comparisons again yielded significant differences in theft rates (all p-values <.025). Consequently, once again only the theft preferences of the elder three age cohorts were subjected to binomial analysis.

Proposer Analysis 3 – did a majority prefer theft in each condition?

The results of the binomial analysis are shown in Figure 3. At 6-7 years, children preferred theft only in conditions in the selfish theft (2/1;1, proportion=.7, p=.017), justifiable theft (1/2;1, proportion=.8, p<.001), modest theft (0/1;3, proportion=.85, p<.001) and zero-sum conditions (0/4;0, proportion=.75, p=.002). However, they preferred non-theft in the pointless theft 2 condition (3/0;1, proportion=.22, p=.001). Unlike Experiment 1, 6-7 year olds in

Experiment 2 were more circumspect regarding obviously unfair offers, demonstrating no outright preference for greedy theft (3/1;0, proportion=.58, p=.43). They also displayed no overall preference in the responder advantage control condition (0/0;4, proportion=.53, p=.875).

In the 8-9 years cohort, this pattern was similar, but with one key difference. Children again refrained from preferring theft in greedy theft (3/1;0, proportion=.48, p=.883) and responder advantage control (0/0;4, proportion=.46, p=.659). However, this time they also no longer demonstrated a preference for selfish theft (2/1;1, proportion=.61, p=.184). They did, however, continue to prefer theft in the justifiable, modest, zero-sum and (all p-values <.002), and to prefer non-theft in the pointless theft condition (3/0;1, proportion=.28, p=.005).

Amongst those aged 10+ years, the pattern was very similar to those aged 8-9 years, another replication of Experiment 1. No preference was recorded for greedy theft (3/1;0, proportion=.46, p=.689) or selfish theft (2/1;1, proportion=.5, p>.99). All other conditions saw a significant preference for theft (all p-values <.001), except in the pointless theft 2 condition (3/1;1, proportion=.16, p<.001), in which subjects preferred not to pull the Theft handle.

Finally, pairwise McNemar tests were performed comparing rates of greedy theft 2 (3/1;0), selfish (2/1;1) and justifiable (1/2;1) theft in order to probe for proposers' ability to distinguish these actions at each stage. At 6-7 years, subjects did not vary their theft rates between any of these conditions (all p-values >.05). By 8-9 years, however, subjects preferred justifiable theft (1/2;1) to selfish theft (2/1;1), p=.002, and also justifiable theft to greedy theft 2 (3/1;0), p<.001, though they did not differentiate between greedy and selfish theft (p=.454). The same pattern was true of the 10+ cohort, who inhibited greedy (p=.005) and selfish (p=.018) theft relative to justifiable theft, but also did not differentiate between greedy and selfish theft (p=.86).

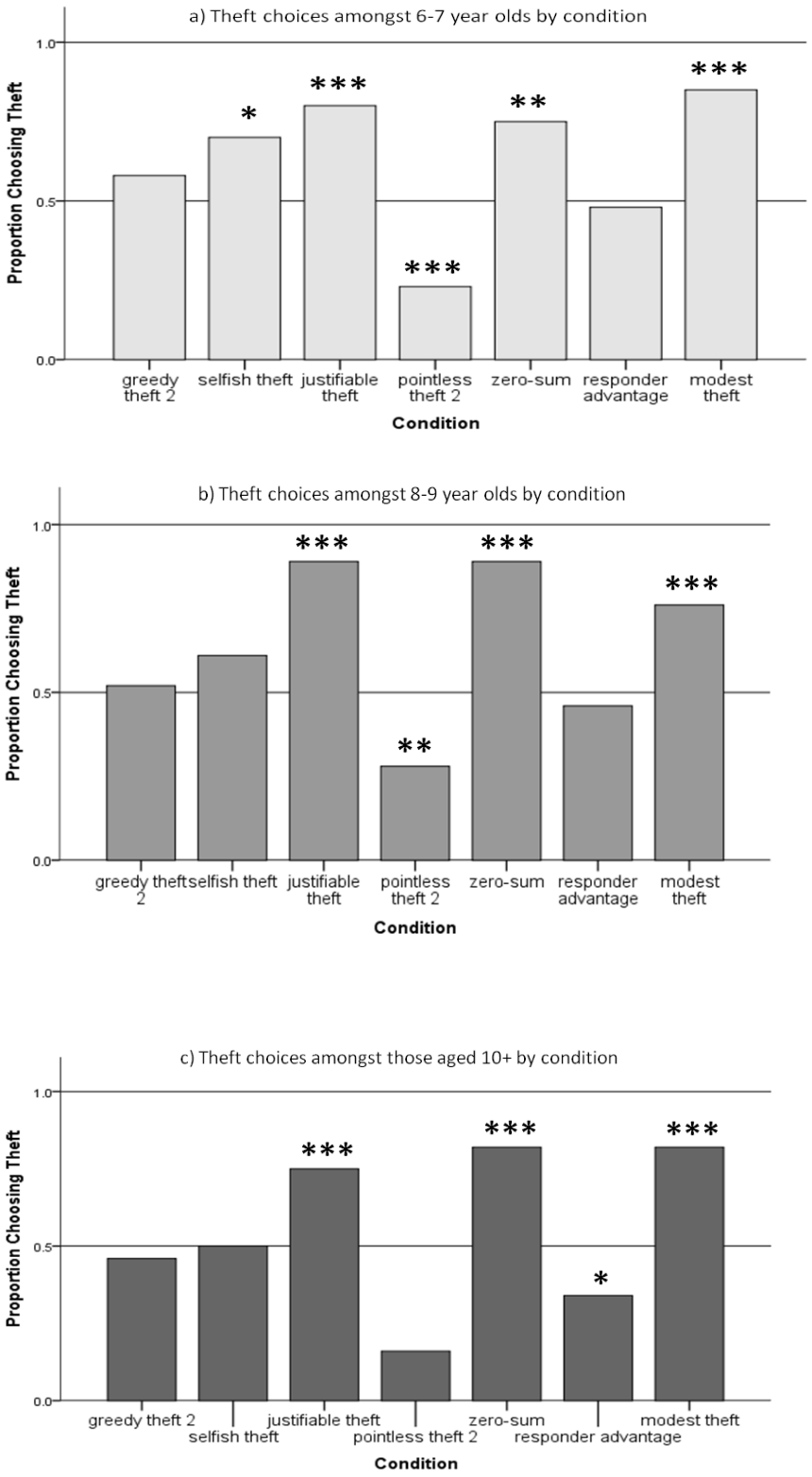

Figure 3: Proposer theft rates in Experiment 2.

*Theft rates are shown by age cohort and condition, with binomial analysis indicating a preference for either theft or non-theft in each case: *p<.05, **p<.01, ***p<.001.*

4.3.2: Responders

All responders successfully passed the knowledge probes, so all tested subjects (n=165) were included in the analysis. Overall rejection rates of the offer caused by proposer theft were as follows: greedy theft 2 (3/1;0) – 69%, justifiable theft (1/2;1) – 50%, pointless theft 2 (3/0;1) – 53%, selfish theft (2/1;1) – 52%, modest theft (0/1;3) – 87%, responder advantage control (0/0;4) – 7%; zero-sum (0/4;0) – 74%. Also of interest in Experiment 2 are the rejection rates of offers of 25% (one point) in response to non-theft, i.e. an unfair outcome which was caused by the experimenter and not the proposer. These rates were 47% in the pointless theft 2 condition (3/0;1), and 54% in the greedy theft condition (3/1;0). A noteworthy finding suggested by these figures is that the one point offer was rejected by approximately 50% of the sample in all five of the contexts in which it was made (i.e. in all of the following: justifiable theft, pointless theft 2, selfish theft, pointless 2 non-theft, greedy 2 non-theft).

Responder Analysis 1 – logistic regression

Stepwise logistic regressions examined whether age in months, gender and relationship predicted rejection of offers *in response to theft only*, i.e. the figures quoted in the previous paragraph. The results of these regression analyses are presented in Table 8. Age in months was an important predictor in six of the seven conditions, generally predicting higher rejection rates – i.e. greater violations of the SEM – with the exception of the responder advantage control condition, in which it predicted fewer rejections of the 4 points offer. There was one gender effect, with boys more likely than girls to reject offers of 3 points in the modest theft condition, although even male rejections were only 20% in this condition. There were no effects of relationship upon rejection rates in Experiment 2.

In a second set of stepwise logistic regressions looking at responses to *non-theft* only, there were two conditions in which age in months positively predicted higher rejections. These were the greedy theft 2 (3/1;0) and pointless theft 2 (3/0;1) conditions, meaning that in both cases older children were once again more likely to reject offers of one point. In all of the other conditions, in which non-theft left the responders with fair or advantageous offers (i.e. in conditions of selfish, justifiable, zero-sum, responder advantage and modest theft), there were no significant demographic predictors of rejection. This means that that participants of all types simply preferred to accept fair and advantageous offers at a uniform rate in these five conditions.

Table 8: significant predictors of offer rejection in Experiment 2.
The table shows significant predictors of rejection in response proposer theft in all conditions, and also to non-theft in two relevant conditions.

proposer choice	*condition*	*predictors*	*β*	*Wald*	*p*
proposer theft	greedy theft 2 (3/;01)	age in months	-0.017	4.92	0.027
	selfish theft (2/1;1)	age in months	-0.017	5.74	0.017
	justifiable theft (1/2;1)	age in months	-0.015	6.06	0.014
	pointless theft 2 (3/0;1)	age in months	-0.033	8.26	0.004
	zero-sum (0/4;0)	age in months	-0.029	14.75	<.001
	res advantage (0/0;4)	age in months	0.088	4,38	0.036
	modest theft (0/1;3)	boys > girls	-2.033	6.86	0.009
proposer non-theft	greedy theft 2 (3/;01)	age in months	-0.017	4.92	0.027
	pointless theft 2 (3/0;1)	age in months	-0.033	8.26	0.004

Responder Analysis 2 – did rejection rates differ from null?

Participants were again grouped into the four age categories (4-5 years, 6-7 years, 8-9 years, 10+ years). Within each condition and within each age category, McNemar's test was used to determine whether the number of rejections differed significantly from null. Data showing rejection rates of theft in each condition at each age group is shown in Table 9.

The data is relatively similar to that from Experiment 1, but there are some important additions. In particular, there is greater evidence of deviation from the SEM. In the new justifiable theft condition, and in this condition only, children of all ages violate the SEM by rejecting offers of one point at a rate significantly greater than null. This includes pre-schoolers (p=.031). In all other cases of 25% offers, however, significant SEM violations again do not occur until 6-7 years, after which point they remain consistent in all conditions at all age groups (see the theft columns for pointless theft 2 and selfish theft, and also the non-theft columns for greedy theft 2 and pointless theft 2). Note that once again the rejection rate of the 100% (four points) offer in the responder advantage control (0/0;4) does not differ from null in any of the age cohorts, in replication of Experiment 1.

Responder Analysis 3 – did subjects favour acceptance or rejection of theft?

As in the previous Experiment, binomial analysis was performed comparing rejection rates in the above data to a baseline of 50%. This was done to see whether a majority of subjects preferred to accept or reject offers in each condition. The results of this binomial analysis are shown in Table 10.

Table 9: responder violations of the SEM in Experiment 2.

*McNemar analyses show whether rejection rates differed from zero within each age cohort: *p<.05, **p<.01, ***p<.001.*

		condition and notation								
		theft							non-theft	
age group	preferences	greedy theft 2 3/1;0	justifiable theft 1/2;1	pointless theft 2 3/0;1	selfish theft 2/1;1	modest theft 0/1;3	responder advantage 0/0;4	zero-sum 0/4;0	greedy theft 2 3/1;0	pointless theft 2 3/0;1
4-5 years	% favouring rejection	38%	23%	8%	16%	7%	19%	35%	40%	28%
	different from null?	yes**	yes*	no	no	no	no	yes**	no	no
6-7 years	% favouring rejection	80%	41%	67%	42%	4%	11%	76%	50%	37%
	different from null?	yes***	yes***	yes**	yes**	no	no	yes***	yes**	yes***
8-9 years	% favouring rejection	84%	64%	90%	71%	21%	0%	88%	72%	63%
	different from null?	yes***	yes***	yes**	yes***	yes*	no	yes***	yes***	yes***
10+ years	% favouring rejection	73%	60%	64%	62%	10%	0%	86%	46%	53%
	different from null?	yes***	yes***	yes**	yes***	no	no	yes***	yes***	yes***

At 4-5 years children did not prefer to reject offers in any condition, even those in which they were left with zero. Instead, in 5 out of the 7 conditions in which they were left with at least part of the endowment (including 25% in selfish theft, justifiable theft and pointless theft 2), a significant majority chose to accept (all p-values <.035). They were also indifferent as to whether to accept offers of zero in the greedy theft and zero-sum conditions, despite the fact that the proposer was left with 100% in both of these conditions. In other words, children of a pre-school age adhered very closely to the SEM at this stage of the analysis, and these results therefore represent a close replication of the findings in Experiment 1.

Table 10: responder rejection preferences in Experiment 2.

*Binomial analyses show whether responders prefer to accept or reject theft by condition and age cohort: *p<.05, **p<.01, ***p<.001.*

		condition and notation								
		theft							non-theft	
age group	preferences	greedy theft 2 3/1;0	justifiable theft 1/2;1	pointless theft 2 3/0;1	selfish theft 2/1;1	modest theft 0/1;3	responder advantage 0/0;4	zero-sum 0/4;0	greedy theft 2 3/1;0	pointless theft 2 3/0;1
4-5 years	% favouring rejection	38%	23%	8%	16%	7%	19%	35%	40%	28%
	different from 50%?	no	accept**	accept**	accept**	accept***	accept*	no	no	no
6-7 years	% favouring rejection	80%	41%	67%	42%	4%	11%	76%	50%	37%
	different from 50%?	reject*	no	no	no	accept***	accept***	reject*	no	no
8-9 years	% favouring rejection	84%	64%	90%	71%	21%	100%	88%	72%	56%
	different from 50%?	reject**	no	reject*	reject*	accept***	accept***	reject***	reject*	no
10+ years	% favouring rejection	73%	60%	64%	62%	10%	100%	86%	46%	53%
	different from 50%?	reject*	no	no	no	accept***	accept***	reject***	no	no

By 6-7 years, a significant majority preferred to reject offers of zero in cases of greedy theft (3/1;0, p=.012), and zero-sum theft (0/4;0, p=.005). The only majorities in favour of acceptance were in cases where they were faced with offers of at least 50% (responder advantage control 0/0;4, p=.001; modest theft 0/1;3, p<.001). This is also a very close replication of Experiment 1.

By 8-9 years, responders retained their majority preferences for accepting offers in both conditions in which they were offered at least 50% (responder advantage control, 0/0;4, p<.001; modest theft, 0/1;3, p=.001). They also retained their preferences for punishing offers of zero (greedy theft 2, 3/1;0, p=.004; zero-sum, 0/4;0, p<.001). At this age majorities preferred to reject offers of 25% in three circumstances. In response to theft, the majority rejected offers of 25% in the pointless theft 2 (3/0;1, p=.021) and selfish theft (2/1;1, p=.036) conditions. In response to non-theft, they rejected offers of 25% in the greedy theft condition (3/1;0, p=.043). These preferential rejections of 25% by 8-9 year olds are the only example in either experiment whereby a majority of subjects was observed to be actively in favour of costly punishment.

By 10+ years responders had reverted to the same patterns as at 6-7 years, preferring to reject offers of zero (greedy theft 2, 3/1;0, p=.029; zero-sum, 0/4;0, p<.001) and accept offers above 50% (responder advantage control, 0/0;4, p<.001; modest theft, 0/1;3, p<.001), but with no strong preference towards offers of 25%.

4.3.3: Comparing Decisions in Experiments 1 and 2

Table 11 presents and contrasts the total number of proposer thefts and responder rejections in each condition in both experiments. As Section 1 of the Table shows, there was a high degree of replication of findings within conditions which remained unchanged from Experiment 1 to Experiment 2. Section 2 shows the divergence in subjects' behaviours between the two experiments in conditions in which the design was altered in Experiment 2.

Table 11: comparison of subjects' preferences in E1 and E2.

The frequency of proposer and responder preferences in the repeated and altered conditions across the two experiments.

		allocation		proposer thefts		responder rejections	
	condition	E1	E2	E1	E2	E1	E2
Section 1: unchanged conditions	selfish theft	2/1;1	2/1;1	59%	58%	32%	38%
	zero-sum	0/4;0	0/4;0	83%	82%	62%	64%
	responder advantage	0/0;4	0/0;4	47%	45%	11%	7%
	modest theft	0/1;3	0/1;3	82%	79%	13%	12%
Section 2: altered conditions	greedy theft	2/2;0	3/1;0	65%	52%	48%	61%
	fair/justifiable theft	1/1;2	1/2;1	77%	79%	17%	41%
	pointless theft	2/0;2	3/0;1	31%	24%	19%	48%

4.4: Discussion

Seven hypotheses were tested in this experiment. Five of these (H1-H5) were targeted at responders, and two of them (H6-H7) targeted proposers. **Hypothesis 1**, that only responders of 6-7 years and above would violate the SEM, was partially supported. All children aged 6-7 and above did indeed consistently violate the SEM. At pre-school level, children were far closer to the SEM although they did deviate from it once. **Hypothesis 2** was supported, in that the older children got the more likely their responses were to violate the SEM by enacting second party punishment. **Hypothesis 3** was supported. Just as in Experiment 1, older children were more likely to enact cost-free as well as costly punishment, suggesting that they are more punitive generally. **Hypothesis 4** was supported. In replication of Experiment 1, there were no age effects in the rejection of fair or advantageous offers, with all age cohorts actively preferring to accept these types of offers. **Hypothesis 5** was not supported. There was little evidence of older children, or indeed any other subjects, differentiating their responses based on a concern with outcome and intention. **Hypothesis 6** was not supported. There were few linear age increases in proposer generosity in Experiment 2. However, this was not due to increased generosity, as **Hypothesis 7** was supported – proposer theft rates were relatively high at all ages, and at no point did proposers actually prefer *not* to steal when doing so could potentially gain them an advantage.

4.4.1: Responder discussion

Generally, the responder findings in Experiment 2 are very consistent with those in Experiment 1, thus adding considerable support to its findings. The key difference between the two experiments from a responder's perspective was that this time there were more conditions (four as opposed to one) in which offers which were both greater than zero and smaller than the proposer's allocation could be made. These were greedy theft 2 (3/1;0), selfish theft (2/1;1), justifiable theft (1/2;1) and pointless theft 2 (3/0;1)[2]. These conditions were the prime candidates for encouraging spiteful punishment, due to the potential for the responder to be faced with disadvantageous inequity.

Offer Rejection Patterns in Experiment 2

Rejection patterns in Experiment 2 were very similar to Experiment 1. Older responders were more likely to enact spiteful punishment of disadvantageous inequity in response to theft, and this occurred in all four of the key conditions listed above. Older children were thus more willing to violate the SEM, and therefore more hypercompetitive. This corresponds with several previous developmental studies (Bereby-Meyer & Fiks, 2013; Blake & McAuliffe, 2011; McAuliffe et al., 2013; McAuliffe et al., 2014). It also provides an interesting mirror to previous work on positive ORPs, which has shown that older children are also more likely than younger children to violate the SEM via pro-social behaviours (Benenson et al., 2007; Fehr et al., 2008; Gummerum et al., 2010; Kogut, 2012; Smith et al., 2013).

Importantly, however, older children were once again also more willing to punish cost-free theft, meaning that their higher punishment levels are not simply due to a greater willingness to tolerate selfish costs. Instead, it seems that they are more prone to punishment behaviour of all types, and that punishing is an action which many pre-school children simply fail to engage in, even when doing so costs them nothing and the punishment is arguably merited. A prime example of the latter would be a proposer making an offer of zero in the greedy theft or zero-sum conditions, a frequent occurrence. This is clearly totally unfair, particularly in the greedy theft condition, and young children should be well aware that this is so (Blake et al., 2014; Smith et al., 2013). Nevertheless, many pre-schoolers seemed largely unmoved by such actions, as once again there was no active preference for punishing this type of behaviour. This is in stark contrast to the other age cohorts, all of whom clearly demonstrated that they

[2] Note that in this condition theft made no difference to the outcome, so the offer was always 25% regardless of the proposer's actions.

considered this behaviour unacceptable via high punishment rates. In further replication of Experiment 1, however, there were no age effects concerning the acceptance of fair or advantageous offers. Responders in all age cohorts considered these to be satisfactory, and consistently expressed strong preferences for accepting them.

Pre-schoolers were, once again, very close to the SEM in their responses to low offers. There was one exception to this, which was a rejection of theft in the justifiable condition at a rate significantly above null. However, caution needs to be expressed about reading too much into this. Jointly, Experiments 1 and 2 involved a lot of pairwise comparisons, thus inflating the risk of a type 1 error. In mitigation of this, the majority of the preferences observed across the two were both replicated very closely on multiple occasions, and characterised by very small p-values (frequently <.001), thus allowing confidence in the results. However, in the case of pre-school children, one deviation from null out of six attempts[3] across two experiments is not convincing evidence of other-regarding preferences. Several recent MUG studies with pre-schoolers have now observed similar behaviours in 4-5 year olds (Bereby-Meyer & Fiks, 2013; Takagishi et al., 2014; Wittig et al., 2013). None of these studies compared pre-schoolers to older children, but my findings from Experiments 1 and 2 suggest that this largely self-regarding behaviour undergoes a big change at around 5-6 years of age, with a strong increase in punitive other-regarding preferences beginning to emerge at around this age. This suggestion chimes well with recent research showing that most children are incapable of appropriately targeting their reciprocity towards deserving recipients until 5.5 years and upwards (House et al., 2013). The development of children's reciprocal behaviours will be explored in depth in the next chapter.

A final important note to add on the topic of responder preferences is that, again in replication of Experiment 1, at no point did the number of subjects rejecting the responder control offer of 100% of the allocation differ significantly from null, meaning that virtually all children of all ages preferred to accept 100% all of the time in both experiments, and that offers of 100% were unique in this respect. The repeated replication of this finding in two experiments with a combined sample of 315 subjects is very good evidence in favour of both task comprehension and a motivation towards accruing prizes, particularly when placed alongside responders' universal success in passing the knowledge probe trials.

[3] In Experiment 1, selfish theft (low offer) only; in Experiment 2, selfish theft (low offer), justifiable theft (low offer), greedy theft (high offer), pointless theft 2 (both high and low offers).

Outcome and Intention Understanding in Experiment 2

The extra conditions in Experiment in which offers of 25% could be made allows for greater insight into whether or not responders took the proposer's intentions into consideration. Responders in Experiment 2 largely failed to make the outcome/intention distinction. Regardless of whether or not the offers of 25% were the product of the more generous proposer choice (greedy theft 2, 3/1;0), the less generous proposer choice (justifiable theft, 1/2;1; selfish theft, 2/1;1), or even when the proposer had no choice at all (pointless theft 2, 3/0;1), the rejection rate always hovered around 50%, and responders of 6-7 years upwards failed to consistently prefer to reject or accept any of them. Instead, they simply appeared divided as to whether 25% is an acceptable offer.

This is, as discussed in Chapter 3, standard adult responder behaviour in a UG with linear offers (Camerer, 2003), but in the MUG most adults tend to be much more forgiving of disadvantageous inequity when it was the best offer available (Falk et al., 2003; Radke, Güroğlu, & De Bruijn, 2012). The lack of difference in rejection rates suggests that subjects in Experiment 2 were attending mostly to simple outcomes. Another interesting thing to note here is that in *all five* of the circumstances in which responders were offered 25% (see Table 8, Section 4.3.2, p.85), older children were more likely than younger children to reject these offers. Note that this means older children were more likely to punish the 25% offer not just in conditions in which it was a product of the proposer's meanness (e.g. selfish theft), but also when it was actually the more generous offer (greedy theft 2), or when the proposer had no choice in the matter (pointless theft 2). When outcome/intention understanding has been previously observed, it has been in children aged around 8-years and over (Bereby-Meyer & Fiks, 2013; Sutter, 2007). In the present study, however, the eldest subjects were simply more likely to violate the SEM by flatly rejecting all unequal offers, regardless of what actually caused them. Thus, although they certainly displayed greater ORPs towards avoiding disadvantageous inequity, they do not appear to have cared what actually caused this inequity to occur. This preoccupation with outcome over intention, the so-called "outcome bias", is something which most humans of all ages suffer from to at least some extent, but it seems that pre-adolescents are particularly prone to this (Castelli et al., 2014).

This finding is consistent with several recent studies concluding that proposer understanding of (or at least willingness to consider) the outcome/intention distinction is not common amongst those in late childhood (Castelli et al., 2014; Gummerum & Chu, 2014; Güroğlu et al., 2009)Thus, I conclude that the ability to move beyond a conception of fairness based primarily

on equity (Fehr & Schmidt, 1999) and relative comparison (Bolton & Ockenfels, 2000) to encompass one that also includes an understanding of others' intentions and realistic capacity to act (Falk et al., 2003, 2008; Nelson, 2002) does not reliably emerge in the majority of people until adolescence. This is probably due to the fact that our perspective-taking abilities do not fully mature until the mid-to-late teenage years (Blakemore, den Ouden, Choudhury, & Frith, 2007; Dumontheil et al., 2010; Gummerum & Chu, 2014; Güroğlu et al., 2009; Radke et al., 2012).

4.4.2: Proposer discussion

The proposer findings represent a slightly less close replication of Experiment 1 than the responder findings. Fewer replications of the linear age effects were observed. Nevertheless, overall theft rates appear to have been very consistent between Experiment 1 and Experiment 2, as Table 11 (see Section 4.4.3, p.87) demonstrates, as indeed does a comparison between Figure 3 (p.82) in this chapter and Figure 2 in Chapter 3 (p.62). The patterns of proposer choices are virtually identical in those conditions which are replicated across the two experiments. From a proposer's perspective, the changes to the experimental conditions in Experiment 2 were also important. The proposer was now presented with a greater number of opportunities for advantageous inequity that did not necessarily require offering the responder nothing. This was now possible in four conditions (selfish theft, 2/1;1; greedy theft 3/1;0; justifiable theft,1/2;1; and pointless theft, 3/0;1), as opposed to a single condition in Experiment 1 (selfish theft, 2/1;1). Unfortunately, another aspect of Experiment 2 which replicated Experiment 1 was that pre-school proposers did not demonstrate an ability to understand the procedure, with many failing the knowledge probes, and the cohort as a whole failing to differentiate the control conditions. The following conclusions therefore apply only to children aged 6-years and above.

In common with Experiment 1, proposers in Experiment 2 were very prone to making clearly unfair offers, and never actually expressed an active preference against this behaviour at any age-group, in contrast to the conclusions of many prominent previous studies (Blake & McAuliffe, 2011; Fehr et al., 2008; House et al., 2013; McAuliffe et al., 2013; McAuliffe et al., 2014; Smith et al., 2013). Once again, these tendencies persisted in a very large proportion of the sample, including those of 8-9 years of age and above. The only point of departure is in the greedy theft condition. Here, children in Experiment 2 were much more likely to refrain from this act at an earlier stage (6-7 years), rather than adopt this restraint as they got older. Given

children's strong known preferences for relative advantage (Blake et al., 2014; Sheskin, Bloom, et al., 2014), it may be that the default offer of 75% in this condition in Experiment 2 (as opposed to 50% in Experiment 1) was enough to invoke greater satisfaction with the default offer, and therefore reduce the impulse towards further theft. Indeed, a likely candidate for the reason behind the absence of proposer theft age effects in Experiment 2 is simply that proposers in Experiment 2 were happier to accept the default allocation, which tended to be advantageous. Certainly this explanation would be consistent with the behaviour of the responders in both experiments, who consistently accepted relative advantage, with no variation in this preference according to age.

Amongst both proposers and responders, then, I observed very little reluctance to accept advantageous inequity. Proposers chose advantageous distributions frequently, and responders preferred to accept advantageous offers in virtually all circumstances. This suggests a distinct lack of a preference for fair and equal outcomes, a theme which was certainly replicated from Experiment 1 to Experiment 2. Comparison of Figures 2 and 3 (p.63; p.83) is again instructive here – note the complete lack of an active preference for not stealing in the greedy and selfish theft conditions at any age group in either experiment (in spite of proposers clearly demonstrating an ability to preferentially choose non-theft in the pointless theft control condition).

Experiment 2, then, successfully replicated the novel finding from Experiment 1 that 8-years of age does not necessarily represent a watershed moment in development at which a concern for fair and equal outcomes comes to characterise children's economic preferences. This is, as discussed, in contrast to many previous findings on this topic (Blake & McAuliffe, 2011; Fehr et al., 2008; House et al., 2013; McAuliffe et al., 2013; McAuliffe et al., 2014; Smith et al., 2013).

It seems likely that the "theft" framing of the MUGWT was the key factor behind this finding, despite the fact that the words "steal" or "theft" were never actually used in the procedure. It could be argued, with considerable justification, that this aspect of the experimental design carried with it certain demand charactersitics in favour of making more anti-social proposals. Nevertheless, I feel it is still legitimate to suggest that the results show the need for more caution when making firm statements in favour of children's preferences for fair outcomes, when these outcomes conflict with their own self-interest. Experiments 1 and 2 show that these can be easily undermined. Real intuitive preferences, however, should be robust to experimental framing cues. The primates tested in previous MUGs, for example, were totally consistent in their decision making regardless of whether the experiment was framed as a

giving or taking exercise (Jensen et al., 2007a; Kaiser et al., 2012). Furthermore, the conventional "giving" or "offer" method of creating distributions in dictator games and ultimatum games has corresponding demand characterstics of its own, which favour the artificial inflation of the generosity of proposer preferences (Bardsley, 2008; Winking & Mizer, 2013). A final important point to make is that although the results on proposer fairness preferences are unusual, they are not unprecedented. For example, Fehr et al (2008) is often cited as a landmark study in unearthing the emergence of egalitarian fairness preferences at around 8-years of age, but it should be noted that a team of researchers from the same laboratory themselves recently failed to replicate this finding in a follow-up study using an almost identical methodology to the 2008 study (Fehr et al., 2013).

4.4.3: Conclusion

Together, Experiments 1 and 2 represent the largest ever ultimatum game study performed on a developmental sample, and also only the second (after Murnighan & Saxon, 1998) to chart other-regarding preferences from the pre-school years all the way through to adolescence using the same experimental procedure. Overall Experiment 2 built successfully on Experiment 1, unearthing useful findings on the fairness rules that children use to guide their decisions, and an encouraging degree of replication as to the pattern these decisions take.

Chapters 3 and 4 have shown that punitive preferences of all types, costly and non-costly, increase throughout childhood, and therefore children punish more with increasing age regardless of whether these actions violate the SEM. From 6-7 years, if not before, humans develop an unusually strong negative other-regarding preference for punishment, including costly punishment. Certainly, it has not been observed in other species in comparable experiments; indeed, our closest genetic relatives adhere perfectly to the SEM as responders in the ultimatum game (Jensen et al., 2007a; Kaiser et al., 2012; Proctor et al., 2013). Experiments 1 and 2 have also shown that children's other-regarding preferences appear to be chiefly motivated by relative comparisons, specifically in the form of preferences for advantageous inequity and against disadvantageous inequity. Across a very broad age spectrum, there was little evidence of the types of fair behaviour that would be expected if other – arguably more sophisticated – fairness concerns were in evidence, such as a consideration of outcome and intention.

Chapter 5: The Moonlighting Game

5.1: Introduction

Chapters 3 and 4 have examined the emergence in ontogeny of second-party punishment, both costly and non-costly. Direct retribution of this kind can be termed "reciprocal" behaviour, in that the responder pays back an actor's behaviour in kind. So, when the actor's opening gambit is positive, reciprocity involves a positive action in return; and when the actor's first move is malign, reciprocity requires retribution. Chapters 3 and 4 looked specifically at the latter which, as the introduction made clear, is an understudied aspect of human behaviour, and particularly of children's behaviour. This chapter aims to extend this work by moving beyond looking at negative reciprocity in isolation, and attempting to relate it quantitatively to positive reciprocity throughout the course of childhood, using a single experiment, in order to allow direct comparison of the development of both reciprocal tendencies. This has rarely, if ever, been attempted before in a published experiment, and indeed there are very few existing examples of experiments suitable for the task, but one which does satisfy all of the necessary criteria is the moonlighting game (Abbink, Irlenbusch, & Renner, 2000), outlined below, which will here be adapted for children for the first time. This chapter will therefore extend the scope of the dissertation by comparing children's reciprocal preferences of both types, thus studying the strength of pro- and anti-social other-regarding preferences together using the same methodology. In doing so, it will also allow exploration of the assumptions underlying strong reciprocity theory (Gintis, 2000b), a prominent evolutionary theory of human sociality.

5.1.1: Aims and Objectives

1) To examine how the strengths of both positive and negative reciprocity change throughout ontogeny.
2) To directly compare the strength of the two motivations, within-subjects, using a single experiment.
3) To explore some of the assumptions underlying Gintis' (2000) strong reciprocity theory.
4) To pioneer the use of the moonlighting game in a developmental sample for the first time.

Reciprocity and Strong Reciprocity

Understanding reciprocity is central to understanding social behaviour. Classic models from evolutionary game theory have long shown that reciprocal strategies, i.e. those which repay kindness with kindness and meanness with meanness, are vital to the maintenance of cooperation over the course of repeated interactions (Axelrod & Hamilton, 1981; Trivers, 1971). Direct reciprocity of this kind is thus considered a necessary lynchpin for cooperative behaviour to emerge and stabilize within a population (Nowak, 2006). To *Homo sapiens*, however, reciprocity is more than simply an evolutionary strategy. It is one of the most fundamental norms of our social behaviour and a cornerstone of human morality (Gouldner, 1960; Haidt & Joseph, 2004). Furthermore, it appears that it may be a human universal to occur to at least some degree in all world cultures (Boehm, 2012), a suggestion supported by global cross-cultural experimental work (Henrich et al., 2001; Henrich et al., 2005; Herrmann et al., 2008). For reasons such as these, the tendency to reciprocate has been suggested as having an innate basis in human cognition (Haidt & Joseph, 2004).

Such findings have led to the suggestion that humans, uniquely in nature, are "strong reciprocators" (Fehr et al., 2002; Gintis, 2000b), meaning that to enforce the reciprocity norm they are willing to pay a cost to either reward or punish and, crucially, they will do this regardless of whether there is any of chance of them reaping net material benefits in the long term.

Positive and Negative Reciprocity – two sides of the same coin?

Strong reciprocity theory tends to assume that both positive and negative reciprocity are driven by the same mechanism[4]. This is an intuitively appealing idea, and one that is implicitly shared by many others among those who seek to explain our violations of the SEM. For example, this is true of two of the theories of the fairness preferences that underlie ORPs, which were outlined in Chapter 1. Both Bolton and Ockenfels (2000) and Charness and Rabin (2002) cite "reciprocity" as a key motivator of ORPs. Theories of social norms are similar in this respect, and here too a fixation with "reciprocity" is conceived of as one of the most basic of our societal rules (Gouldner, 1960; Haidt & Joseph, 2004).

4 For example: "A strong reciprocator is predisposed to cooperate with others and punish non-cooperators, even when this behaviour cannot be justified in terms of extended kinship or reciprocal altruism" (Gintis, 2000, p.169).

However, there is a catch. Previous attempts to experimentally test the theory that positive and negative reciprocity are mirrored expressions of one's concern with fairness have tended to find little evidence of a correlation between the two types of reciprocal tendency within the same individual (Keysar, Converse, Wang, & Epley, 2008; Peysakhovich, Nowak, & Rand, 2014; Yamagishi et al., 2012). Personality psychology studies using questionnaire methods have also tended to find the absence of any significant within-subjects correlation between metrics purporting to measure positive and negative reciprocity as separate constructs (Dohmen, Falk, Huffman, & Sunde, 2008; Perugini et al., 2003), although these studies suffer somewhat from a tendency to implicitly frame negative reciprocity in a pejorative manner in the questions they pose to participants.

Regardless of which of these theories is correct, one important point to make about reciprocity is that it frequently causes systematic violations of the SEM (Bolton & Ockenfels, 2000; Charness & Rabin, 2002; Gintis, 2000b). Positive reciprocity does so in a way that is hyper-cooperative, and negative reciprocity does so in a way that is hyper-competitive, but both involve paying an additional cost, targeted towards another individual, in response to an event that is already in the past. In both cases, the "rational" thing to do is simply to walk away and not spend any more of one's time and resources responding to the initial event. It is true that repeated interactions between actors can create scenarios in which reciprocal behaviour is rational instead of other-regarding (Axelrod & Hamilton, 1981; Trivers, 1971). However, theories of strong reciprocity, fairness preferences and social norms are all clear on the fact their predictions apply just as much to scenarios in which reciprocity is made for its own sake, and where there is no prospect of future gain. Repeated findings from diverse economic experiments show that these assumptions are largely correct (Camerer, 2003).

Children's Positive and Negative Reciprocity

When considering children's reciprocal tendencies, the existing literature appears to suggest something of an imbalance between positive and negative reciprocity. Chapters 3 and 4 have already reviewed a wealth of literature on children's ultimatum game behaviour as responders. This tends to show that, at least from middle childhood upwards, many children are more than willing to reciprocate negatively by spitefully rejecting low offers (Bereby-Meyer & Fiks, 2013; Castelli et al., 2014; Gummerum & Chu, 2014; Güroğlu et al., 2009; Murnighan & Saxon, 1998; Sutter, 2007). However, the same cannot necessarily be said of positive reciprocity. In behavioural economics, the classic test of positive reciprocity is the trust game (Berg, Dickhaut, & McCabe, 1995). In the trust game, a first player chooses how much of an

endowment to offer to a second. The amount offered is then tripled by the experimenter, and the second player is asked if they would like to return any of the endowment to the first (i.e. positive reciprocity). Children, on the whole, are not especially trustworthy as the second player in this experiment, and more of them elect to keep the endowment without offering any back to the first player than is generally the case with adults (Evans, Athenstaedt, & Krueger, 2013; Sutter & Kocher, 2007). The finding that younger children are less positively reciprocal than older children, and that both are less so than adults, was also found in a conceptually similar study using an iterated prisoner's dilemma (Gummerum et al., 2009).

The UG and trust game are, of course, different experiments, so comparing the levels of reciprocity shown in each is not necessarily a sound guide to the strength of children's preferences. This is particularly true in light of the demonstration in the previous two chapters that children's behaviour does not always generalise particularly well from one UG study to the next, let alone to a completely different experiment. In the sphere of reciprocity, however, there are few other sources to look to, since very little has been done in the way of attempts to directly compare positive and negative reciprocity in childhood. One notable exception to this is Keil (1986), who found that negative reciprocity was stronger in childhood than positive reciprocity. However, a problem with this experiment was that the two types of reciprocity did not involve equivalent costs. The positive reciprocity required a violation of the SEM, and was thus hyper-cooperative, but the "negative reciprocity" was actually profitable, and thus rational in the classical economic sense. The fact it was more common is thus easy to explain via more parsimonious mechanisms than reciprocity. Another recent study to explore this issue is that of van den Bos et al., 2012, who also found that negative reciprocity outstripped positive reciprocity in childhood. Here again, however, the two types of reciprocity did not involve equivalent costs, since "positive reciprocity" was essentially a risky action in which the payoff could vary substantially, whereas negative reciprocity generated a medium, intermittent payoff. Isolating exactly what caused the different responses to each type of event is therefore difficult. Nevertheless, the findings of studies such as these can form the basis of testable hypotheses.

The Moonlighting Game

To satisfy the criteria of strong reciprocity theory, both types of reciprocity need to be costly and thus systematically violate the SEM. One of the few existing behavioural economics experiments which, firstly, includes the potential for positive and negative reciprocity and, secondly, is structured so that both involve paying an equal cost is the moonlighting game

(Abbink et al., 2000). It is this experiment upon which this chapter will focus, and a full explanation will be given in the Methods section, below. Briefly, however, the experiment involves two players (who in the interests of simplicity I will continue to refer to as "proposer" and "responder"[5]), who are initially presented with equal endowments. The proposer moves first, and decides to either steal part of the responder's endowment, or gift the responder part of their own endowment. If the proposer steals, the responder can reciprocate negatively by paying to eliminate the proposer's endowment at a ratio of 3:1, in other words costly punishment. If the proposer gives, their gift is multiplied by a factor of three, and the responder has the option of gifting back however much of this windfall they like in return, in other words costly positive reciprocity. This positive reciprocity subgame is directly analogous to the trust game (Berg et al., 1995).

In the original moonlighting game, adult responders employed a high frequency of spiteful punishment in the negative reciprocity subgame, but were less consistent in their reciprocation in the positive reciprocity subgame. Nevertheless, the heavy cost of punishment meant that the positive reciprocity subgame was still the more likely of the two to prove profitable. Interestingly, adult proposers largely appeared to anticipate this and preferentially chose the prosocial subgame (Abbink et al., 2000). It is worth noting, however, that a subsequent version of the experiment, this time using the strategy method, found much less variation between levels of positive and negative reciprocity amongst responders (Falk et al., 2008). It is therefore difficult to draw firm conclusions as to whether or not adults display different levels of the two types of reciprocity within-subjects, though this would certainly be a worthy topic of future research.

Focus of the Present Chapter

How children behave in this experiment is unknown, although the work reviewed above allows some inferences to be drawn. It seems likely that negative reciprocity will exceed positive reciprocity in pre-adolescents (Keil, 1986; van den Bos, van Dijk, & Crone, 2012). The ontogeny of the two tendencies is less easy to predict, but based on the findings of the previous two chapters, as well as those from much of the existing children's behavioural economics literature already reviewed, it seems reasonable to assume that older children will be more likely to deviate from the SEM, and thus show greater levels of both types of reciprocity.

[5] Abbink et al (2000) simply refer to them as "player A" and "player B".

Another focus of this chapter is the closely related question of whether positive and negative reciprocity are correlated within-subjects, as strong reciprocity theory tends to assume. Although previous work with adults has suggested that they are not (Dohmen et al., 2008; Keysar et al., 2008; Perugini et al., 2003; Peysakhovich et al., 2014; Yamagishi et al., 2012), none of these studies has actually compared the two tendencies directly and quantitatively within the same experiment, as will be the case here. Finally, proposer behaviour is also difficult to predict. Chapters 3 and 4 have shown proposers in a not dissimilar scenario to the present experiment behaving in a manner which was both greedy and largely un-strategic. Furthermore, the cited developmental literature on the trust game (Evans et al., 2013; Sutter & Kocher, 2007; van den Bos et al., 2012) has shown child subjects to be more reluctant to make trusting overtures than adults. Both of these findings suggest that proposers may be more inclined towards theft than giving, in spite of the fact that the latter is potentially the more profitable decision. This notwithstanding, previous research has found that trust increases with age during development (Evans et al., 2013; Sutter & Kocher, 2007) alongside other prosocial sentiments such as altruism (Fehr et al., 2013), whereas anti-social sentiments such as envy decline (Fehr et al., 2013; Steinbeis & Singer, 2013). For this reason, I tentatively predict that gifts will increase in frequency and size in conjunction with age, and the reverse patterns will be seen for thefts.

5.1.2: Hypotheses

Five hypotheses will be tested in this chapter. The first three (H1-H3) are based on previous findings on patterns of reciprocal behaviour in children. The final two (H4-H5) are aimed at exploring strong reciprocity theory:

- **(H1)** Negative reciprocity will increase over the course of ontogeny.
- **(H2)** Positive reciprocity will also increase over the course of ontogeny.
- **(H3)** Negatively reciprocal behaviours will exceed those for positive reciprocity.
- **(H4)** The two types of reciprocity will show a positive relationship within-subjects, i.e. those who reciprocate positively will also be more likely to do so negatively, and vice-versa.
- **(H5)** In common with Experiments 1 and 2, proposer behaviour will tend towards advantageous inequity over disadvantageous inequity, and thus be characterised by more taking than giving.

5.2: Method

5.2.1: Participants

Three hundred children aged 6-12 years (mean age = 9.12 years, SD = 2.06 years; 144 girls/156 boys) were recruited and tested at the Live Science Area is the usual manner. The minimum age of 6 years was selected for two reasons. The first is that previous findings have shown that this is the age at which most children become capable of targeting positive reciprocity at those who have helped them, and negative reciprocity at those who have not (House et al., 2013). The second is a more pragmatic one – pre-school proposers in the MUGWT did not reliably demonstrate an ability to inhibit actions of theft, and it was felt that this might act as a confound in the moonlighting game, which used a similar apparatus with a theft component.

All subjects volunteered, with the written consent of their parents, to take part in the experiment at the Live Science Area, over the course of 8 weekends. Children were tested in pairs, and were matched for age as closely as possible. The mean disparity in age between the members of each pair was 1.8 years (SD = 1.35).

5.2.2: Design

In common with the MUGWT studies of Chapters 3 and 4, the design involved a between-subjects division into "proposer" and "responder" roles. Within these roles, all subjects were subjected to four counterbalanced habituation trials (see Table 12, p.104), and then six identical experimental trials. Unlike Chapters 2 and 3, there were not multiple conditions. Instead, there was a single task which was repeated over six trials. Subjects were randomly allocated to the role of either proposer or responder, and stayed in this role for all six trials. There were therefore 150 proposers (mean age = 9.13 years, SD = 2.18; 71 girls, 79 boys) and 150 responders (mean age = 9.11 years, SD = 1.95; 73 girls, 77 boys). Within each pair, both participants took part in four habituation trials and six experimental trials in their allotted role.

Proposers' and responders' responses were analysed separately. For proposers, the outcome variable of interest was whether they preferred actions of giving or theft, and the strength of each of these preferences, as measured by the number of points they gave or stole on each

trial. For responders, the key outcome variables were how frequently and strongly they reciprocated acts of giving or theft with costly acts of their own. These consisted of giving or punishment respectively.

In both cases, three between-subjects factors were included in this analysis: age in months, gender, participants' relationship with their partner (binary-coded: pre-acquainted, or "known" (1), versus unacquainted, or "unknown" (0)). Participants were habituated face-to-face and simultaneously in the same pairing in which they undertook the experiment, and using the same apparatus. All 300 participants were subject to the same four habituation trials prior to the commencement of the experiment.

5.2.3: Materials

Trials were conducted using the apparatus shown in Figure 4. The apparatus consisted of a square, white-painted, wooden board, approximately 1m x 1m. On top of this board were placed two smaller boards, one marked "giving" and the other marked "taking", in large, red writing. Both consisted of the same two elements: three ramps, which were fixed to only tip in one direction, and one cylindrical tube, fixed to tip down from the responder's end towards the proposer. On the board marked "giving", the three ramps were fixed to tip away from the proposer, towards the responder. On the board marked "taking", the ramps were fixed to tip away from the responder towards the proposer. Down the centre of all three of the boards was painted a large, green stripe, clearly dividing the board into two equally-sized halves, one each for both proposer and responder.

The points used with the apparatus were once again exchanged post-experiment for prizes in the manner explained in Chapter 2. During the experiment, each participant's accumulation of points was stored in a transparent plastic tub placed next to them. The proportions of points allocated to proposer and responder on each trial were the same as in the original moonlighting game (Abbink et al., 2000), but the numbers were halved in the interests of simplicity, i.e. the initial endowment of 12 became six, the maximum prospective gift/theft of six was halved to become three, etc.

a) **b)**

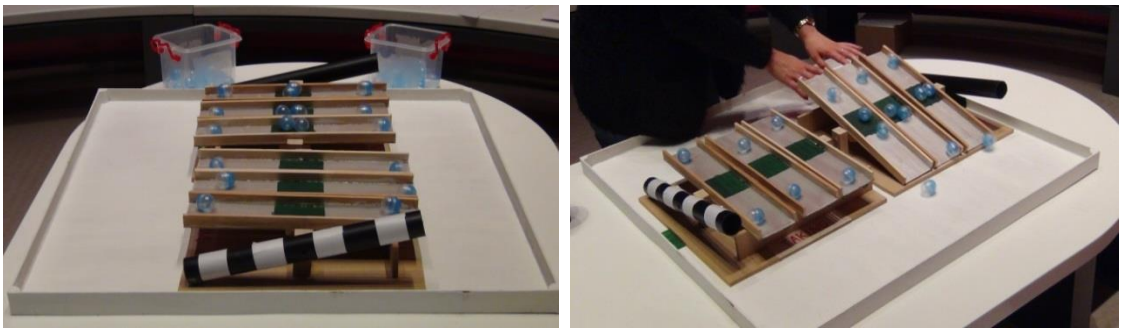

Figure 4: The apparatus in Experiment 3 before and during play.
In a) the board is arranged in the standard initial distribution common to all trials, with 6 points on each side of the green line, and 6 additional points in the centre of the giving board. In b) the proposer is giving points to the responder, by tipping the ramps on the giving board. Doing so causes not only the 3 points on the proposer's side to move across, but also the 6 points in the middle. The proposer is thus left with 3 points (6-3), and the responder with 15 (6+3 from the responder, + the additional 6 from the central neutral zone).

5.2.4: Procedure

Participants were first introduced to the experiment's prize-structure. They were shown the array of prizes and told the number of necessary points needed to acquire each prize. All participants were told that they needed "to fill up your bucket as much as you can". Once this was done, in an attempt to emphasize the (potentially) positive- rather than zero-sum nature of the experiment's prize structure, it was stressed to participants that it was the fullness of their own transparent bucket that would win them prizes, and that the contents of their partner's bucket would *not* affect their own prize.

All pairs were then presented with the same four counterbalanced habituation trials (see Table 12). In all cases, as in the experiment proper, the apparatus was initially calibrated by the experimenter in the distribution shown in Figure 4, with six points placed on each side of the green line, one at each tip of all of the ramps. In addition, on *only* the board marked "giving", six additional points were placed in neutral territory on the green line in the middle of the giving ramps, with two at the centre of each ramp.

On all habituation trials, it was stressed to children that every point on their side of the line belonged to them, and those in the middle initially belonged to nobody. To reinforce this, children were asked on all trials, both during habituation and the experiment, to verbally count out how many points were sat in front of them at the start of each trial (always six). They were

then asked to confirm how many points were in the middle (six), and state who the points in the middle belonged to ("nobody").

Table 12: sequential explanation of the habituation trials in E3.
The table summarises the four habituation trials presented to all participants, and the permutations within each. Each habituation trial is represented by one column. Each stage of each habituation trial is explained row by row. In all trials both proposer ("prop") and responder ("res") begin with six points by default.

stage	habituation trial			
	moderate prosocial	generous prosocial	moderate antisocial	strong antisocial
1) proposer's action	giving board; 1 ramp	giving board; 3 ramps	taking board; 1 ramp	taking board; 3 ramps
2) result of prop's action	prop 5; res 7	prop 3; res 15	prop 7; res 5	prop 9; res 3
3) responder's action	return 2	return 6	burn 1	burn 3
4) result of res's action	prop 7; res 7	prop 9; res 9	prop 4; res 4	prop 0; res 0
5) net profit =	both +1	both +4	both -2	both -6

Habituation Trials – Proposer Phase

During every habituation trial it was explained to children that any given trial could only involve *one* of the sub-game boards on the table ("a giving round or a taking round"), and that in the experiment it would be the proposer's job to decide which was used. On each of the habituation trials, the proposer was asked to alter the balance of the initial distribution to the relevant degree, depending on which scenario was being demonstrated (see Table 12). Broadly, each individual act of giving would confer three extra points upon the responder (the proposer's one, plus two from the neutral zone), whereas each act of taking would simply move one point from the responder's side over to the proposer. This was demonstrated as follows. When the ramps on the giving board were tipped, this caused the point on the proposer's end of the ramp and, additionally, the two points in the neutral zone on each ramp's central green line, to move across to the responder's side. This effectively multiplied each gift from the proposer by a factor of three. When the taking board was tipped, the only alteration of the balance was the simple movement of one point per ramp away from the responder's side, over towards the proposer.

Habituation Trials – Responder Phase

After the initial balance had been altered, both children would be asked to count aloud how many points were now on their side. They would then be asked how many they started with (i.e. six), in order to reinforce the manner in which the balance had changed in, or against, their favour. The responder would then be asked to return the appropriate number of points, depending on the trial (see Table 12) down the tube on the active board. On the *giving* board, this action (using "the giving ramp") returned points back to the proposer for banking in their transparent bucket. On the *taking* board, this action (using "the burning ramp", marked differently, see Figure 4) "burned" the points on the proposer's side at a ratio of 3:1. In practice this involved the experimenter physically removing the appropriate number of points from the proposer's side of the board. Again, as in the case of giving trials, children were asked to count aloud how many points they each had left, and then to compare this to how many they started with (six). All points remaining at the end of this process could then be banked in the same manner as with giving trials. Children would do this by physically picking up the points in front of them, and placing them in their transparent prize bucket.

Final Instructions

At the end of the four habituation trials, the prize structure and absolute rather than relative nature of the scoring was reiterated, with particular emphasis. Children were reminded that the other player's score did not affect their prize. It was stressed that they simply needed not to worry about the other player's score, and just fill up their own bucket with points as fast as they could. At this point, children's prize buckets were emptied of all of the points accumulated on the habituation trials, and it was explained that the process was now about to begin "for real." Both participants were then told that the experimenter was no longer going to tell them what to do: from now on it was "up to you".

Experimental Trials

All six experimental trials took the following format. The board was set up in the usual manner and children were first asked to once again repeat to the experimenter how many points were set up on their side of the board (six), and how many points were in the middle (six). All participants were able to answer these questions without hesitation by this stage.

On each trial, the proposer was asked two questions: firstly, did they want this to be a "giving" or "taking" round? Once this was answered verbally, the experimenter asked them to tip as many or as few ramps as they liked on the appropriate board only (i.e. from 1-3, depending

upon their preference). At this point, participants were once again asked to count aloud how many points were on their side of the line. On the rare occasions when participants made a mistake, they were politely asked to try again, until they gave the correct answer. They were then asked, yet again, how many points they had been allocated at the start of the trial (six). This procedure was designed to reinforce the sense of loss or gain. Again, participants did not struggle to promptly give the correct answer (six).

The responder was then told it was their turn. They were asked how many points they would like to roll down the giving or burning tube respectively, dependent upon the proposer's initial choice of board. In the case of the giving board, they were told that they could return as many or as few as they liked, for the proposer to keep. In the case of the taking board, they were told they could spend from 0-3 points, in order to burn the proposer's points. Responders were first asked for a verbal answer. Once they had given this, they were told "ok, go ahead," and allowed to physically carry out their request. In rare cases where responders requested to change their mind and execute a different physical action to their verbal request, they were allowed to do so, and the physical number of points transferred was the score coded as their decision for the trial. Once the process of responder reciprocation had been completed, participants were once again asked to count how many points they had left, and, one final time, how many they had begun the trial with (six). Once the correct answers had been given, they were told to put their points "in the bank". On this command, participants would pick up the remaining points on their side of the green line, and place them in their transparent prize bucket.

At the end of the six experimental trials, the total number of points in each participant's prize bucket was counted up, and participants were allowed to choose their prize accordingly. Participants and parents were then thanked and debriefed. The total procedure took around 15-20 minutes.

5.3: Results

The same division of analysis between proposer and responder results as in Chapters 3 and 4 was employed. Both the proposer and responder analyses are again broken down into two subsections. The "First Trial Data" subsection was analysed first and separately, as on this trial observations of participants' decisions were truly independent, i.e. unaffected by the other player's preferences, which at this stage were still unknown. The "Overall Data" subsection

looked at pooled data across the six trials, focusing on the aggregate behaviour of each participant over the course of repeated interactions with their partner/opponent. In all sections of the analysis the standard demographic predictors (age in months, gender, relationship, see Chapter 2) were included in all regression equations. The results section concludes with a look at the prevalence of different types of reciprocity from responders, in order to see if any of them exhibit the patterns of behaviour which strong reciprocity theory would predict.

5.3.1: Proposers

First Trial Data

Eighty four participants (56%) made the decision to steal on the first trial, and thus choose the anti- over the pro-social subgame. Binomial analysis showed this majority to be non-significant ($p=.165$). A stepwise logistic regression analysis looking at give/take as a binary outcome using the standard four predictors found that none of them significantly predicted this decision. Although the frequency of pro- and anti-social choices was not significantly different on trial 1, the magnitude of thefts, in terms of the mean number of points, was larger. When choosing theft, proposers took a significantly higher number of points for themselves ($M=2.46$, $SD=.69$) than they were willing to give away when making gifts ($M=1.53$, $SD=.73$; Mann-Whitney, $U=1103$, $p<.001$).

Overall Data

Across all trials, acts of theft were substantially more common than acts of giving. Out of the combined total of 900 trials (150 proposers x 6 trials), 59.22% involved thefts and 40.33% involved gifts. A full breakdown of the frequency of each type of theft or gift is shown in Table 13. As shown by section a) of the table, the modal proposer decision out of the six possible actions was the maximum possible theft (take 3). Furthermore, looking at gift trials only, the most common choice was the least generous gift (1).

The sum total of decisions by each proposer to either give (1) or take (0) across all trials was calculated, thus generating a binary gift/theft index total score between 0 and 6. Had there been no preference for either action, this score would therefore have been 3. Instead,

participants significantly preferred theft, generating a mean score of 2.43 (SD=1.07, t(149)=-6.5, p<.001).

Table 13: frequency of a) proposals, and b) responses, in E3.
Percentages of each proposer choice, and of the responder reciprocations to each type of choice, across all 900 trials (i.e. 150 participants x 6 trials).

a) proposer's offer		b) responder's reciprocation							
		return 3+	return 2	return 1	return 0	burn 0	burn 1	burn 2	burn 3
offer type	%=	%=	%=	%=	%=	%=	%=	%=	%=
gift 1	24.74	8.6	16.3	14.5	60.6	-	-	-	-
gift 2	8.84	20.3	20.3	11.4	48.1	-	-	-	-
gift 3	7.05	42.9	3.2	6.3	47.6	-	-	-	-
take 1	8.06	-	-	-	-	36.1	20.8	33.3	9.7
take 2	16.35	-	-	-	-	37.7	28.1	19.2	15.1
take 3	34.94	-	-	-	-	33.3	23.7	14.7	28.2

As Table 13 suggests, as well as being more common, acts of theft (mean size in points = 2.4, SD=.55) were also, within-subjects, larger on average than acts of giving (M=1.53, SD=.53, Wilcoxon, Z=-9.21, p<.001). Two separate stepwise regressions were performed with the demographic variables as predictors, the first with mean gift size as the outcome variable and the second with mean theft size as the outcome variable. No models containing any of these variables were found to predict mean theft size. However, one significant model was identified for mean gift size (R^2=.032, (F(136, 1)=4.45, p=.035). This model contained only age in months but found that increasing age in months predicted smaller gift-sizes (B=-.004, t=-2.127, p=.035). Finally, within-subjects Spearman correlation analysis of mean theft and mean gift size revealed no significant linear relationship between these two variables (two-tailed, n=144[6], r_s(142)=.14, p=.103).

The proportion of binary decisions to give or take was observed to vary over the course of the combined six trials (Cochran's test, Q(5)=11.34, p=.045). A comparison of the first and final trials showed that acts of giving had declined by the end of the experiment relative to the beginning (McNemar, X^2=5.49, p=.019). Given this fall in giving behaviour over the course of iteration, binomial analysis was performed in order to examine at which point proposers' preferences for taking or giving actions became significant. There were no preferences for giving, but on the final two trials was there a statistical preference for acts of theft (trial 5 observed prop=.65, p<.001; trial 6 observed prop=.68, p<.001). This means that a preference for taking rather than giving had emerged in proposers' decisions in the later trials of the experiment.

[6] N=144, and not 150, due to six proposers from the original 150 who always stole and thus failed to generate a giving score.

5.3.2: Responders

First Trial Data

The observed frequency of negative reciprocation to thefts, at 55 out of 84, was significantly higher than the observed frequency of positive reciprocation to gifts at 32 out of 66, as demonstrated by chi-square analysis ($\chi^2(1)=4.38$, p=.036).

Overall Data

A full descriptive breakdown of responders' reciprocations to each proposer choice is given in Table 13 (above). Two scores were generated for each responder based on the proportion of times they elected to reciprocate both types of proposer act. In other words, all responders were awarded one score for their proportion of positively reciprocal responses, and a second score for their proportion of negatively reciprocal responses. Within, subjects, the proportion of negatively reciprocal responses (M=.64, SD=.35) was larger than the proportion of positively reciprocal responses (M=.39, SD=.38) to a significant degree (Wilcoxon, Z=-5.09, p<.001). However, the modal reciprocity choice from responders in both cases was zero (for negative reciprocity, zero occurred in 34.7% of responses; for positive reciprocity zero occurred in 55.6% of responses).

Next to be examined was how the proposers' decisions affected the *size* of the responders' positive and negative reciprocity, i.e. the number of points the responder was willing to pay to reciprocate in each case. Two stepwise regressions were performed, with mean punishment size and mean reciprocal gift size respectively as the outcome variables. In both regressions, the standard demographic predictors were included. In the case of positive reciprocity, proposer's mean gift size was included as an additional predictor. In the case of negative reciprocity, proposer's mean theft size was included as an additional predictor. For the giving regression, mean gift size only (b=.552, t(136)=3.2, p=.002) was found to predict positive reciprocity in the only model to achieve significance (R^2=.07, F(1, 137)=10.26, p=.002). For the negative reciprocity regression, however, no combination of predictors produced a significant model, and furthermore an additional block-entry regression was nowhere near significant (R^2=.03, F(5, 133)=.83, p=.533). The absence of a positive relationship between age and reciprocity was contrary to predictions in both cases.

The potential for a relationship between the tendencies towards both positive and negative reciprocity was examined. This was done in order to probe some of the assumptions of strong reciprocity theory. First, Wilcoxon analysis was used to compare within-subjects the mean number of points payed by responders to reciprocate thefts and gifts respectively. Negative reciprocity payments (M=1.29, SD=.93) were significantly larger than those for positive reciprocity (M=.91, SD=1.1; Wilcoxon, Z=-3.65, p<.001).

This, however, was to be expected. The provocation for negative reciprocity was generally larger (i.e. because thefts tended to be bigger than gifts), so it is no surprise that the response was too. A more direct comparison between the two types of reciprocity was attempted by binary coding responses to each proposer act as either balanced (1; equal to, or bigger than, the proposer's act) or weak (0; smaller than the proposer's act). This allowed each responder to be assigned a mean score for their proportion of balanced acts of positive reciprocity (M=.35, SD=.36) and negative reciprocity (M=.38, SD=.36). Within-subjects Wilcoxon analysis showed there to be no significant difference between these scores (Z=-.603, p=.55). Also notable is the fact that there was no significant within-subjects correlation between subjects' proportionate positive and negative reciprocity scores (Spearman, two-tailed, $r_s(145)=.079$, p=.34).

Additionally, there was no within-subjects correlation between positive and negative reciprocity in terms of mean points spent on each outcome (Spearman, two-tailed, n=144, $r_s(142)=.047$, p=.56). Finally, there was no relationship between the frequency of decisions to respond to both gifts and thefts (Spearman, two-tailed, n=144, $r_s(142)=.072$, p=.39). Together, these results suggest no simple relationship between levels of positive and negative reciprocity within-subjects.

5.3.3: Sub-types

This section aims to provide an overview of attempts to find consistent strategic behaviours amongst proposers and responders.

Proposers

Proposers were classified, based on their overall pattern of choices, into one of the following categories: strongly pro-social (always giving: 6 gifts/0 thefts), mostly pro-social (mostly giving: 4-5 gifts/1-2 thefts), neutral (even use of both strategies: 3 gifts/3 thefts), mostly anti-social (mostly taking: 1-2 gifts/4-5 thefts), strongly anti-social (always taking: 0 gifts/6 thefts). Few proposers adopted either of the extreme strategies, with only 4% electing the strongly anti-social strategy, and not a single subject (0%) adopting the strongly prosocial strategy. Of the intermediate strategies, mostly anti-social was the modal outcome (51%), ahead of neutral (31%) and mostly pro-social (15%).

Responders

Initial classification attempts searched for two specific subtypes. Firstly, those who could be catalogued as *strong reciprocators*, i.e. those who "cooperate with others and punish non-cooperators" (Gintis, 2000, p.169). The second involved those who adhered to the SEM and consistently refused reciprocity of both kinds, thus demonstrating a lack of other-regarding preferences. People of this type are often referred to in the literature as "selfish" (Falk et al., 2008; Fehr & Fischbacher, 2002; Levine, 1998). Both of these strategically pure subtypes were largely absent from the sample. Only 7% of subjects adopted a strong reciprocity strategy, and only 6% could be classified as selfish.

These results show that although blanket reciprocity is not a common strategy, neither is zero reciprocity – instead, participants appeared to be adopting more subtle decision rules. Consequently, a second attempt at responder categorization was made using a more moderate selection of strategic choices, listed in Table 14, which required consistency of response to only *one* type of reciprocity, as opposed to both. In order to show "consistency" in response to proposer behaviour, it was deemed necessary for responders to be exposed to a behaviour at least twice, meaning that those who were only exposed to zero or one instance of proposer gifts/thefts were excluded from the analysis. This left a remaining *n* of 128. Note that these moderate strategies are *not* all mutually exclusive (with the exception of the no strategy pattern). For example, it is entirely possible for an uncharitable responder's actions to qualify them as someone who never rewards generosity, and yet always punishes theft. It is for this reason that the *n* and percentage columns in Table 14 add up to greater than 100%.
Less than 20% of the sample (here labelled "no strategy") was seen to show no pattern at all to their behaviour. All other responders showed consistency in at least one of the reciprocal strategies listed. In terms of the focus of this chapter, the most striking result in the table is the fact that the percentage of responders who chose to negatively reciprocate all anti-social acts

perpetrated against them is almost exactly double that of those who chose to always positively reciprocate generous behaviour.

Table 14: responders' decision rules in Experiment 3.

strategy	n=	%=
always reward	23	17.97
always punish	46	35.93
never reward	46	35.93
never punish	26	20.31
inconsistent	25	19.53

5.4: Discussion

Five hypotheses were tested in this experiment. **Hypothesis one**, that negative reciprocity would increase in conjunction with age, was not supported. **Hypothesis two**, that positive reciprocity would increase in conjunction with age, was also not supported. **Hypotheses three** was supported: negatively reciprocal behaviour did exceed positively reciprocal behaviour. **Hypothesis four** was not supported. There was no positive relationship, or indeed any relationship, between positive and negative reciprocal tendencies within-subjects. **Hypothesis five** was supported. Proposers preferred the anti-social taking sub-game to the pro-social giving sub-game.

5.4.1: Comparing Positive and Negative Reciprocity

There was mixed support for the hypotheses concerning negative reciprocity. The most important success, in terms of the predictions made, concerns the fact that negative reciprocity exceeded positive reciprocity, a finding which was in line with expectations. A simple and striking example of this is provided by the fact that responders reciprocated negatively to a majority of thefts, but only chose to reciprocate positively to just over a third of gifts. Previous applications of behavioural economics experiments to developmental samples have hinted that a tendency for costly negative reciprocity may be more common in childhood than positive reciprocity (Bereby-Meyer & Fiks, 2013; Evans et al., 2013; Gummerum et al., 2009; Sutter & Kocher, 2007; van den Bos et al., 2012). However, this work involved testing one or other of these preferences separately, whereas the present experiment tested them

together in a comparable manner. The findings show that pre-adolescent children have stronger preferences for punishment than they do for acts of reciprocal kindness. All pre-adolescents were therefore more likely to violate the SEM for punitive reasons than they were for reasons of positive reward. This is a very important finding, showing as it does how the mere presence of a tendency to pay a cost to reward should not necessarily be taken as evidence that *Homo sapiens* are inherently benevolent and hyper-cooperative. Such findings need to be placed in context. The present experiment appears to show that our negative other-regarding preferences for spiteful punishment become entrenched in behaviour at an earlier point in development than do positive other-regarding preferences.

One aspect of the findings in this chapter which did not meet the predictions was the lack of variation in levels of negative reciprocity in conjunction with increasing age, a pattern which was also true of positive reciprocity. This is in contrast to the previous MUGWT studies. It is true that Chapters 3 and 4 found the most important leap in negative reciprocity to occur between pre-school (4-5 years) and early middle childhood (6-7 years), and only the latter demographic featured in the present experiment. It is also true that, as outlined in Chapter 1, it is easy to point to prominent experiments in which a linear age effect of increasing ORPs has not been observed (e.g. House et al, 2013; Lergetporer et al., 2014; Murnighan & Saxon, 1998; Sally & Hill, 2006; Sutter, 2007; van den Bos et al., 2012, etc.). A full consideration of age effects in developmental game theory experiments will be made in the discussion in Chapter 7, once all of this dissertation's experimental data have been presented.

5.4.2: Strong Reciprocity

Strong reciprocity was found not to be a prevalent strategy amongst the sample, but equally neither was adhering to the SEM. Instead, a complex collection of individual decision rules along dimensions of both positive and negative reciprocity meant that the developmental sample was showing the beginnings of a tendency towards strongly reciprocal preferences at an *aggregate* level, but not at the individual level. By this I mean that a proposer's chances of having either type of offer reciprocated in a way which violated the SEM were reasonably high, particularly in the case of punishment, since reciprocators of both types were clearly present in the sample. However, the two types of reciprocity were not related within-subjects. Finding that one's partner was a good reciprocal rewarder was not a relevant guide as to whether they would also be a strict punisher.

The exploratory work on decision rules illustrates this, showing that only very small minorities of subjects responded equally to both types of reciprocity by choosing either all or none of the opportunities to respond to both gifts and thefts. Instead, 80% of them appeared to apply separate, more nuanced decision rules to reward and punishment respectively. In conjunction with the increasing body of other, similar findings from experiments indirectly testing this idea (Dohmen et al., 2008; Keil, 1986; Keysar et al., 2008; Perugini et al., 2003; Peysakhovich et al., 2014; Yamagishi et al., 2012), I conclude that "strong reciprocator" is not a common personality type. Instead, an aggregation of varied individual fairness preferences and norms towards both reward and punishment amongst society at large means that humans are strongly reciprocal at a population rather than an individual level. One important caveat to note here is that, strictly speaking, Gintis's (2000) theory of strong reciprocity applies to one-shot rather than repeated encounters. It is true that this experiment did not perfectly model this aspect of the theory , as actors' decisions may have been influenced by the strategic consideration of forthcoming trials. Nevertheless, I still feel that it is striking how little relationship there was between positive and negative reciprocity within-subjects and, in conjunction with the other empirical evidence cited above, I feel that my conclusion is justified.

An interesting implication of this finding is that positive and negative other-regarding preferences are not necessarily products of the same mechanism, e.g. a generic pre-occupation with "fairness" which prompts people to reward do-gooders and punish wrong-doers. This is important, since it is not just strong reciprocity theory which implies that this should be so, but also at least two very notable models of fairness preferences (Bolton & Ockenfels, 2000; Charness & Rabin, 2002), and also prominent work from the norm-psychology literature (Haidt & Joseph, 2004). Instead, it seems that the personality psychologists may be right, and grouping tit-for-tat behaviour of all types under a single label of "reciprocity" may be an over-simplification (Perugini et al., 2003). Much more work is needed to explore this idea, however. Current personality measures of positive and negative reciprocity do not measure the two constructs in a legitimately comparable way, and of the few experimental studies which attempt to compare the two types of reciprocity, several do not assign equivalent economic costs to the two actions (Keil, 1986; Keysar et al., 2008; van den Bos et al., 2012). The moonlighting game does, however, and more exploratory work should be carried out using this under-utilised experiment.

5.4.3: Proposer Actions

Proposers' behaviour was in line with predictions. Thefts exceeded gifts and, just as in the mini-ultimatum game with theft in the previous two chapters, children of all ages showed no reluctance to indulge in acts of theft. This occurred in spite of the fact that the procedure's habituation trials (see Table 12, section 5.2.4, p.104) collectively gave a big hint that acts of reciprocal giving could lead to much more profitable outcomes than acts of taking and punishment. Proposers nevertheless preferred taking behaviour, and this preference become more entrenched on the later trials.

There are two likely reasons why they did this: inequity preferences and lack of trust. Children's fairness preferences surrounding inequity have already been documented in Chapters 3 and 4, i.e. a combination of disadvantageous inequity aversion and, in many subjects, a preference for advantageous inequity. Pre-adolescent children are very reluctant to allow disadvantageous inequity, as the previous chapters have demonstrated, along with many other studies (Bereby-Meyer & Fiks, 2013; Blake & McAuliffe, 2011; Bügelmayer & Spiess, 2014; McAuliffe et al., 2013; McAuliffe et al., 2014; Sheskin, Bloom, et al., 2014). It seems that this aversion may have made them reluctant to upset the initially equitable balance by handing some of their points over to the responder, especially when a tempting alternative was present which allowed them to tip the balance in their own favour instead. One obvious counter to this is that this need only have been temporary, and they could have made a profit in the long run. Here, however, there is a second issue, namely trust. Previous developmental work with the trust game, which is basically just the moonlighting game's giving sub-game in isolation, has shown two things. Firstly, children do not trust each other to reciprocate positively; secondly they are quite right not to do so, since their peers are indeed generally untrustworthy (Evans et al., 2013; Gummerum et al., 2009; Sutter & Kocher, 2007; van den Bos et al., 2012), a finding which was replicated by the present experiment. Thus, if proposers guessed (probably correctly in most cases) that the responder would not reciprocate positively, then the knowledge that their giving would therefore have led to disadvantageous inequity on the trial would have stopped many of them from giving.

There was an additional interesting finding in the proposer data. Increasing age in months negatively predicted the mean amount given away by proposers. This means that this experiment found a rare example of behaviour becoming significantly *less* pro-social in older children. As Chapter 1 made clear, an active decrease in prosociality in older children is a much more unusual finding than the absence of a linear increase.

The most likely explanation for this is probably a strategic one. In a conventional moonlighting game (Abbink et al., 2000), the prediction of the SEM is for the proposer to take as much as possible, on the induction that a rational responder will not want to spend anything on costly negative reciprocity. In the present study, this latter tendency demonstrably failed to occur. Whilst there were high levels of theft, the majority of these events (65%) incurred retribution, thus tending to render them non-profitable. Gifts were also largely unprofitable, with only 16 from 144 proposers (11%) earning net interest on the amount given to proposers, and 56% of transfers going totally un-reciprocated. However, unlike thefts, with gifts the proposer's remaining post-transfer endowment was at least free from the spectre of punishment, and guaranteed to remain un-tampered with. Thus, faced, as many of our proposers were, with an obstreperous responder who refused to reward gifts and yet often acted to avenge thefts, the proposer decision which had the greatest probability of providing the least unprofitable payoff was simply to grudgingly give away as little as possible, whilst hoping for the occasional act of kindness in return. This is a reasonably complex inductive calculation, however, and likely one that older proposers found easier to make. Strategic thinking makes big strides during middle to late childhood (Steinbeis et al., 2012), which is to be expected, since during this stage children's abilities markedly increase in all manner of relevant cognitive domains, such as theory of mind, working memory and impulse inhibition (Reyna, 2012). These are all vital antecedents of making sound strategic decisions in game theoretical contexts (Wischniewski et al., 2009).

Finally it is worth noting how the proposer findings in this chapter relate to the work in Chapters 3 and 4. Unlike the tempting theft handle on the apparatus used in those chapters, in the moonlighting game both pro- and anti-social proposer decisions were enacted by virtually identical mechanisms. Proposers still preferentially chose the anti-social act, suggesting that the findings in Chapters 3 and 4 cannot simply be explained away as the product of demand characteristics attributable solely to the apparatus. When proposers were presented with both choices in a single apparatus in a way that was both visually and mechanistically almost identical, their preference was still for the anti-social act.

5.4.4: Conclusion

This chapter has described the first developmental psychology experiment to compare the strength of costly positive and negative reciprocity in a directly comparable way. Its key finding

is that children are more reliable punishers than they are rewarders. When making reciprocal decisions, pre-adolescent, school-aged children are consistently more likely to violate the SEM in a manner that is hyper-competitive than one which is hyper-cooperative. A second important finding is that at an individual level children's tendency to reward is not related to their tendency to punish, thus suggesting that these other-regarding preferences are distinct, and not simply mirror-image tendencies derived from a meta-concern with a single fairness preference for all acts of reciprocity. In conjunction with similar findings in adults, this suggests that this statement is true of people of all ages, and that hyper-competitive preferences are cognitively distinct from hyper-cooperative preferences.

Chapter 6: The Battle of the Sexes Game

6.1: Introduction

Previous chapters have looked at costly negatively reciprocal behaviour, and how it compares to costly positively reciprocal behaviour, in order to successfully chart the emergence and development of hyper-competitiveness in childhood. This chapter aims to broaden the understanding of hyper-competiveness further by examining how it can affect children's willingness to cooperate in a potentially less adversarial task. The aim is to move beyond looking at hyper-cooperativeness in the narrow context of comparing and responding to different bargaining offers, and instead examine how it can affect children's abilities to reach mutually beneficial, cooperative outcomes. The experiment used in this chapter, the iterated battle of the sexes game (BOS), is one which has potentially greater potential for cooperation than the other experiments used so far. If players are willing to compromise and take turns, mutually beneficial outcomes can be achieved.

A central tenet of this dissertation is that humans are both hyper-cooperative *and* hyper-competitive, yet most existing work focuses only on the former. In this chapter children are presented with an experiment in which they should, rationally, cooperate. Hyper-competitive preferences should thus not be expected to be frequently expressed for either strategic or other-regarding reasons, so whether or not they are will be of considerable interest. In the BOS, a paradigm borrowed directly from game theory, the most mutually beneficial long-term outcome requires one or both players to regularly accept disadvantageous inequity in the short term. For higher joint scores to be achieved, one or both of the players must therefore be willing to overcome the temptation to frame partner interactions in zero-sum terms, and choose cooperative short-term behaviours in the interests of achieving long-term success.

This chapter will apply the BOS to a developmental sample for the first time. This experiment has been chosen as it is one in which cooperation should be favoured by participants according to the SEM. Any deviations from cooperative behaviour are thus potential evidence of hyper-competitiveness. Despite its potential for positive-sum gains, the BOS is not without its pitfalls. Players' outcomes are not independent of their partner's, and yet they are also frequently different from their partner's and thus a degree of social comparison is inevitably involved. Children are highly prone to evaluating their outcomes relative to their peers (Blake et al.,

2014), therefore this is a factor which could promote hyper-competitive preferences. Whether or not children can overcome these urges in order to achieve mutually beneficial outcomes will be of considerable interest.

6.1.1: Aims and Objectives

1) To test whether children's aversion to disadvantageous inequity is strong enough to prevent them cooperating in an experiment in which they can mutually benefit from doing so.
2) To test whether the level of information children have regarding their opponents' intentions will affect their preferences.
3) To examine children's sensitivity to risk, and how it affects their cooperative decision making, by varying the degree to which their cooperative actions can potentially backfire.
4) To examine how encountering a cooperative dilemma over multiple rounds affects children's long-term decision making.

6.1.2: The Battle of the Sexes Game: a cooperative task with elements of social comparison

The BOS is a paradigm borrowed directly from game theory. It involves a scenario in which cooperation can be profitable in an absolute sense, and yet leave the subject exposed to disadvantageous inequity in a relative sense. Examples of the two BOS payoff matrices employed in the present experiment are shown in Table 15. As the table shows, on a single trial, both players are presented with a binary choice between cooperation and defection. If both players defect, they both score one point. If both cooperate, they both score one point (in the Low Risk game) or zero points (in the High Risk game). However, the most profitable outcome is for one to defect and one to cooperate, in which case five points are distributed, split 3-2 in the defector's favour. In other words, as long as one player is prepared to accept being worse off in relative terms (by scoring two as opposed to the other's three), both players can be better off in absolute terms (by scoring two or three instead of one each).

For both players to do well in the BOS, they need to be willing to alternate cooperate and defect choices, since doing so can increase long-term income to a level above what can be

achieved by constantly choosing mutual defection. In other words, by choosing the former they alternately receive outcomes of two or three, as opposed to consistently receiving one.

Table 15: payoff matrices in Experiment 4.

The Low Risk game is used in the Regular condition. The High Risk game is used in the Risk and Sequential conditions.

(A) Low Risk

		P2	
		Cooperate	Defect
P1	Cooperate	1, 1	2, 3
	Defect	3, 2	1, 1

(B) High Risk

		P2	
		Cooperate	Defect
P1	Cooperate	0, 0	2, 3
	Defect	3, 2	1, 1

When both players converge on this strategy, it therefore represents an example of mutualistic cooperation, a term which applies to a cooperative act which benefits all involved. An important fact to note about mutualistic cooperation is that it does not violate the SEM. In a mutualistic scenario, all parties can cooperate for self-serving reasons, and no ORPs are necessary (Baumard, André, & Sperber, 2013; McLoone & Smead, 2014; Tomasello, Melis, Tennie, Wyman, & Herrmann, 2012a). A second important point to note is that even when faced with a consistently defecting partner, a player's most rational decision, according to the SEM, is to consistently *cooperate* in the BOS. This is because in this scenario cooperating confers a score of two on every round, as opposed to one if the player chooses to repay defection with defection. Consequently, explaining cooperation in this experiment is easy – the more interesting behaviour from a psychological point of view is non-cooperation, which in mutualistic scenarios represents a hyper-competitive violation of the SEM.

Indeed, if a player knows that their opponent has defected, or even if they infer or predict that their opponent is going to defect, then choosing to defect in turn is a spiteful decision, since it deliberately incurs a lower payoff for no other reason than to eliminate the relative advantage of one's partner. The experimental design will test this scenario explicitly by including a "Sequential" condition, in which the players take it in turns to make their moves, and contrasting it with two other conditions ("Risk" and "Regular") in which their choices are made simultaneously. If players set aside hyper-cooperative motivations then overall scoring should be much higher in the Sequential condition, since it allows the second mover to easily maximise joint income. If scores in the Sequential condition are the same or lower than in the two simultaneous choice conditions, it will show that the hyper-competitive tendency for zero-sum evaluation is sufficient to overcome adherence to the SEM, and to disrupt mutualistic

cooperation. For a full explanation of all three conditions, see the Design section (Section 6.2.3).

6.1.3: Two factors which could undermine cooperation: inequity aversion and uncertainty

Whilst the BOS is a potentially mutualistic scenario, it is not difficult to conceive of reasons why children may potentially not interpret it as such. There are two factors in particular which may not work in its favour: inequity aversion and uncertainty. Disadvantageous inequity aversion (henceforth "DIA") has already shown itself to be a major motivator of hyper-competitive preferences in this dissertation (see Chapters 3-5), and in many previous experiments, particularly amongst subjects in early to early-middle childhood (Blake & McAuliffe, 2011; Bügelmayer & Spiess, 2014; Fehr et al., 2008; Fehr et al., 2013; McAuliffe et al., 2013; McAuliffe et al., 2014; Sheskin, Bloom, et al., 2014). It is easy to see how DIA could contribute to high defection levels in the BOS. Cooperating may be beneficial in absolute terms, but in the BOS it also means frequently accepting that one's partner will be better off in relative terms. It seems likely that many subjects, particularly at the younger end of the sample, will be unwilling to tolerate this outcome.

Secondly, there is an element of uncertainty in the BOS which is not present for responders in bargaining games such as the ultimatum game and moonlighting. In the BOS, the actor does not know the opponent's intentions. Consequently, choosing defect, with possible outcomes of three or one, may seem more tempting than choosing cooperate, with possible outcomes of two or one. Realistically, one's opponent cannot always be expected to cooperate, but it is possible that wishful or egocentric thinking, combined with a lack of advanced perspective taking skills, may prevent younger children in particular from inferring this. To explore this possibility, different conditions will be employed, manipulating the degree of information and risk present in the game.

6.1.4: Children and Game Theory 1: Other 2x2 Dilemmas

Whilst presenting children with the BOS is a novelty, there are several previous studies which have explored their preferences in other, similar 2x2 game theoretical contexts. These can, therefore, shed some light on what might be expected in the present experiment. These

studies have tended to focus on the prisoner's dilemma (PD) and stag hunt (SH) (see Table 16). In the prisoner's dilemma both players can score well by mutually cooperating (2-2), but the individual's best outcome is always to defect regardless of what their opponent does. This therefore creates a strong temptation for both players to defect, even though mutual defection actually leads to a worse outcome (1- 1) than mutual cooperation. Children, in line with their general tendency to follow the SEM closely, show very low levels of cooperation in the PD when it is played as a one-off encounter. One recent study found defection levels in a large sample of over 1,100 children aged between 7-11 years to be around 80% in a baseline PD (Lergetporer et al., 2014). Findings as to whether or not this tendency changes towards greater cooperation with age in late childhood are inconsistent (Angerer et al., 2015; Lergetporer et al., 2014).

Table 16: payoff matrices for (A) the PD, and (B) the SH.

(A) Prisoner's Dilemma

		P2	
		Cooperate	Defect
P1	Cooperate	2, 2	0, 3
	Defect	3, 0	1, 1

(B) Stag Hunt

		P2	
		Cooperate	Defect
P1	Cooperate	3, 3	0, 1
	Defect	1, 0	1, 1

Such behaviour is again suggestive of limited hyper-cooperativeness in pre-adolescents. However, it must be acknowledged that the PD is not an easy environment for cooperation to flourish. A more benign setting for cooperation's emergence is provided by the SH, in which mutual cooperation is the best possible outcome (3-3), and therefore one which can be favoured for selfish as well as altruistic reasons (Baumard et al., 2013; McLoone & Smead, 2014; Tomasello et al., 2012a). The SH will be focused on in much greater depth in Chapter 7, but for now it is sufficient to say that the developmental evidence to date has found that even in their pre-school years children are skilled cooperators in the SH, succeeding on achieving mutual cooperation on the vast majority of trials (Duguid, Wyman, Bullinger, Herfurth-Majstorovic, & Tomasello, 2014; Wyman, Rakoczy, & Tomasello, 2013). Indeed, by 4-years of age they have already been observed out-performing chimpanzees at this task, due to their superior ability to coordinate a joint behaviour through verbal communication (Duguid et al., 2014). Chimpanzees, it should be recalled, almost always adhere to the SEM and thus should be expected to cooperate in a stag hunt, which indeed they often do (Bullinger, Wyman, Melis, & Tomasello, 2011; Duguid et al., 2014).

The limited evidence to date thus shows that children prefer high defection during the PD and high cooperation during the SH. They will therefore cooperate when doing so is the most potentially beneficial outcome, but not when it is only intermediately profitable and carries a high degree of attendant risk. Both of these responses are rational, according to the SEM, at least insofar as one-shot versions of the experiments are concerned.

6.1.5: Children and Game Theory 2: Iterated Play

There is a further novelty which this experiment will add to the literature, however, which is in going beyond one-shot play and allowing children to contest multiple, repeated trials ("rounds") of the experiment (30 rounds in all conditions). This is common practice in theoretical game theory and in experiments with adults, and rightly so, since it drastically alters the dynamics of play. Even a superficially zero-sum experiment like the PD can result in stable mutual cooperation over repeated rounds, since the knowledge that there are future interactions to come gives both players an incentive to settle on mutual cooperation (2-2) rather than mutual defection (1-1). Disrupting this cooperative equilibrium by defecting can earn more in the short term (3) but is likely counterproductive in the long term as it tends to mean the cooperative equilibrium is replaced by one of mutual defection (Axelrod & Hamilton, 1981). A key thing to note here is that the greater potential for mutualism created by iterated play further reduces the need for ORPs to induce cooperation. The more strategic environment encourages more rational consideration of one's options than is seen in one-shot experiments, with subjects typically paying less attention to factors such as altruistic motives or inequity aversion, and instead basing their play upon a strategic appraisal of the pay-offs (Dreber, Fudenberg, & Rand, 2014). This is important, because it means that deviations from the SEM for other-regarding reasons should therefore be comparatively rare in such contexts, even compared to what would be expected in regular economic experiments.

Experimental evidence as to whether children actually adhere to this prediction is rather thin on the ground, although one important study on children's behaviour in the PD observed that pre-adolescents largely failed to settle on mutual cooperation even in an iterated game (Sally & Hill, 2006). The only other study I know of which subjected children to repeated rounds of an economic experiment involved, unusually, the trust game. Here it was found that 10 year olds' long term cooperative patterns were consistent, but also affected by their opponent's decisions. For example, when faced with a non-cooperative opponent their cooperation levels were consistently low, and when faced with a cooperative opponent they were consistently

high (van den Bos et al., 2012). This, of course, is rational behaviour, and suggests that at least towards the upper end of my age sample, children can target their long-term strategy appropriately, should they be presented with a game and an opponent which present sufficient opportunity. The BOS is, however, obviously a very different experiment to both of these previous examples, and furthermore in both of these studies children were paired with a computer or confederate whose response were pre-programmed, as opposed to with a real partner. It is therefore unknown how pairs of children perform in iterated cooperative dilemmas, and it will be interesting to see whether they can ignore distracting hyper-competitive motivations and achieve lasting cooperative outcomes.

6.1.6: Hyper-competitiveness versus the SEM

Based on the previous findings of other studies from both developmental psychology and game theory, it is possible to make a case for either high or low levels of cooperation occurring in this experiment. Children possess all of the necessary cognitive and social skills for simple cooperative acts by the time they begin formal schooling, and have frequently been observed to be capable of collaborating with partners in cooperative acts during their pre-school years (Tomasello, 2014; Tomasello et al., 2012a; Warneken & Tomasello, 2014). Additionally, although there are only a handful of published studies to date examining children's abilities to coordinate their actions to achieve mutualistic outcomes in pairs, they have all found that this is something that a majority can achieve by 4-5 years of age (Duguid et al., 2014; Grueneisen, Wyman, & Tomasello, 2015; Wyman et al., 2013). Finally, in early-to-middle childhood at least, children are competent rational maximisers (Bereby-Meyer & Fiks, 2013; Takagishi et al., 2014; Wittig et al., 2013). All of these factors should aid them in attaining the cooperative outcome in the BOS. Since, however, most children are also highly attuned to evaluate their lot in relative, zero-sum terms (Bügelmayer & Spiess, 2014; Sheskin, Bloom, et al., 2014) and strongly averse to disadvantageous inequity (Blake & McAuliffe, 2011; McAuliffe et al., 2013; McAuliffe et al., 2014), it remains an open question as to whether they will be able to overcome these conflicting hyper-competitive motivations sufficiently well to achieve high and sustained levels of cooperation in the BOS.

6.1.7: Hypotheses

The work in this chapter is more exploratory than in those previous, since there is less experimental precedent for this study. Nevertheless, given the focus of the dissertation, the following hypotheses are offered, examining hyper-competitive preferences:

- **(H1)** Older dyads will be able to achieve high levels of cooperation in the BOS, but younger dyads will not, due to hyper-competitive concerns with DIA.
- Children will adjust their cooperation levels according to factors of information and risk. They will therefore be **(H2)** more likely to cooperate in the Sequential condition, and **(H3)** less likely to cooperate in the Risk condition, when each is contrasted with the other, and with the Regular condition.
- **(H4)** Iterated play will have a positive effect on the levels of cooperation amongst those in later childhood and their cooperation will be higher in later rounds, but it will have no effect on the cooperation levels amongst younger children, which will be consistently low.

6.2: Methods

6.2.1: Participants

Three hundred and four children aged 5-13 years (mean age = 8.2 years, SD = 1.99 years; 170 girls, 134 boys) volunteered to participate at the Science Museum London's Live Science area. Children sat the experiment in "dyads", meaning pairs which made simultaneous choices on each trial to produce a joint payoff. Care was taken to match dyads for age as closely as possible. The mean age differential across the entire sample was 1.54 years (SD = 1.36 years). This did not vary significantly by condition (Kruskal-Wallis test, $\chi^2(2) = .5$, $p = .78$).

6.2.2: Materials

Decisions to cooperate or defect were indicated by playing cards. All subjects were issued with two decks of cards, one marked with foxes and one marked with lions. The two participants in each dyad were sat at opposite ends of a desk. A sturdy wooden occluder was placed in front of each, in order to hide their un-played cards from one another's view. As in previous chapters, the "points" accumulated were empty gumball capsules and these could be exchanged for prizes at the end of the experiment, as explained in Chapter 2. During the

experiment, each participant's points were placed into a transparent bucket adjacent to where they sat. See Figure 5 for a visual representation of the experimental set-up on a typical trial.

1) Participant's eye view. 2) Bird's eye view.

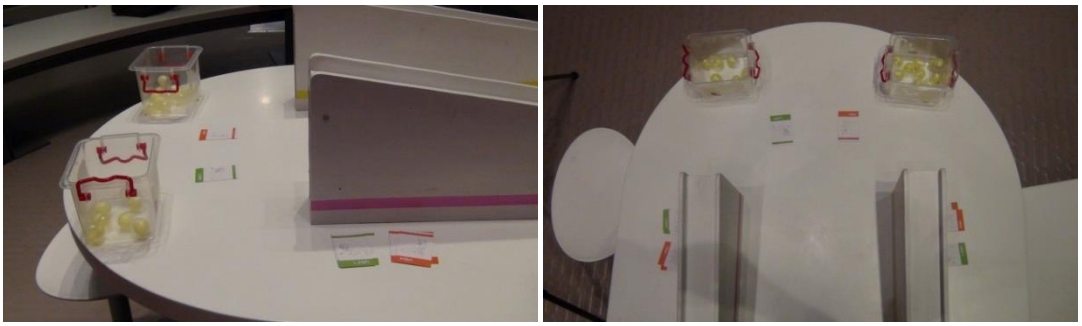

Figure 5: Apparatus used in Experiment 4.
The BOS apparatus. In 1) the game is seen from a participant's eye view. The occluder in the foreground obscures the two players' cards from one another, but allows a clear view of the card table and prize buckets. Un-played cards are in two decks on the participant's side of the occluder, and cards in play are paired in the middle. Image 2) shows a bird's eye view of the apparatus. Each participant has an identical occluder and set of cards in front of them at the bottom of the table, and can only view the playing surface and prize buckets placed at the top. Two unmatched cards are in play.

6.2.3: Design

Experimental Conditions and Payoffs

Dyads were assigned, between-subjects, to one of three conditions. For payoffs in each of these conditions, see Table 15 (p.119). The Regular condition featured conventional BOS payoffs (the Low Risk game) and simultaneous decisions from subjects. The Risk condition featured a BOS with payoffs adjusted to increase the riskiness of subjects cooperating (the High Risk game), and also required simultaneous decisions from subjects. The Sequential condition used the High Risk game payoffs, but required participants to take it in turns to reveal their choices. This meant that the second player to move on each round had full knowledge of how their choice would affect both their own and their partner's payoffs. In all three conditions, 30 rounds of the BOS were played.

As Table 15 shows, all conditions were designed so that they featured two outcomes which were higher in absolute terms, but unequal, (i.e. 2-3 or 3-2). There were either one or two additional outcomes, depending upon the condition. In the Regular condition (Low Risk game), both mutual cooperation and mutual defection resulted in a lower and equal payoff (1-1). In

the Risk and Sequential conditions (High Risk game), mutual defection caused an outcome of 1-1, and mutual cooperation an outcome of 0-0.

Outcomes, Dependent Variables and Scoring

Within each dyad, subjects were assigned a different preferred animal, i.e. one was assigned the fox and the other assigned the lion. "Defecting" involved playing a card showing one's preferred animal. "Cooperating" involved playing a card showing one's partner's preferred animal. The highest scoring outcome in absolute terms (2-3/3-2) involved instances where one partner cooperated and the other defected. In the context of the BOS, this meant both partners choosing *the same card*. This scenario is termed a Match, and the frequency of Matches is one of this experiment's three key dependent variables. By way of illustration, imagine a scenario where the Fox player chooses their own card (Fox) and the Lion player chooses the other player's card (also Fox). The cards are the same, so here we have a Match. However, the Fox player has defected, and thus scores three, whereas the Lion player has cooperated, and thus only scores two (see Table 15, p.119).

A second possible outcome was a payoff of 1-1. In the Regular condition, this was caused by any failure to coordinate the chosen animals. Any mismatched pair of cards (either Fox/Lion or Lion/Fox) therefore resulted in a mutual 1-1 payoff. In the Risk and Sequential conditions, the 1- 1 outcome occurred only in response to *mutual defection.* This second key dependent variable occurred when the Fox player chose Fox and the Lion player chose Lion and was termed a "Clash".

Additionally, the Risk and Sequential conditions also involved the possibility of an additional 0, 0 payoff. This occurred in the case of mutual cooperation, i.e. when both players chose *each other's cards* (i.e. the Fox player chose Lion and the Lion player chose Fox). This third key dependent variable was deemed an Opposite. This least desirable scenario was a genuine hazard in the Risk condition, which involved simultaneous play and therefore guesswork. However, in the Sequential condition Opposites should have been easily avoidable due to turn-taking in playing cards. Thus, the Sequential condition also acted as a control by allowing the checking of task comprehension. A failure to avoid Opposites with a very high degree of consistency in the Sequential condition, when doing so should have been trivially easy, would suggest either a lack of comprehension or motivation on the part of subjects.

6.2.4: Procedure

In all 3 conditions, dyads played 30 rounds of the same BOS, and on each round they were free to choose their own or their partner's preferred card. Dyads were sat facing one another across a table and were each issued with 60 playing cards, split evenly into 2 decks of 30 fox cards and 30 lion cards. In any one session, one deck would represent each player's preference. Each participant had a wooden occluder in front of them which obscured their card decks from their partner's view until they had made their choice. The playing surface on which choices were made was clearly visible to both players. The scoring system involving the use and public distribution of gumballs as pseudo-currency was the same as in all other chapters (see Chapter 2).

Participants in each dyad sat opposite one another, both equipped with identical decks of fox and lion cards. These were placed within easy reach, but obscured from their partner's vision via the wooden occluder. Each participant had a picture of "their" animal pinned to the occluder in front of them, in order to remind them which card represented their preferred animal.

It was explained to participants that they were about to take part in an experiment, in which they needed to try and get as many points as possible in order to exchange them for prizes. At this point, it was made clear that it was their absolute and not their relative score which would determine their final prize allocation. Participants were told: "it *does not matter* how many points the other player gets, ok? The other player's score *does not affect your prize*[7]. All you need to do is to try and fill up your own bucket as fast as possible". Additional steps were also taken to avoid explicitly framing the experiment as a zero-sum game. Words such as "win", "player" and "game" were deliberately omitted from the script.

Next, participants were walked through a habituation phase involving a series of eight demonstration trials, presented in a counterbalanced, pseudo-random order. In the first block of demonstration trials, each of the four potential outcomes in the appropriate experimental condition was demonstrated by asking participants to play the appropriate card in the centre of the table. The appropriate number of gumball capsules (points) was counted out into each participant's bucket during each example. In the Regular and Risk conditions, participants were asked to play their cards face down and then flip them over once both players had made a

[7] In all of the examples quoted in this section, the emphasis is included in the original script.

choice, just as would be the case on the experimental trials. In the Sequential conditions, participants were told to play their cards face up, as would be the case on the experimental trials. In all conditions, the same sequence was then repeated in a second block of four habituation trials (i.e. each of the four outcomes was demonstrated twice). After the second block of demonstration trials, participants were asked to explain the different outcomes back to the experimenter. In the rare event that they failed to give correct answers or hesitated, the demonstration trials were repeated until participants could confidently recite the experiment's payoff structure.

Once the experimenter had ensured that both participants understood the scoring system, both prize buckets were emptied of all points accumulated during the demonstration trials. Participants were told that it was now time to do the experiment "for real", and that all points accrued from now on would count towards their final prize. At this point, immediately prior to the commencement of the experimental trials, the fact that the experiment's scoring system was relative and not absolute was once again heavily emphasized to participants in all conditions. Participants were told: "remember the very important thing I told you earlier. The other person's score *does not affect your prize*. To get a prize, you just need to fill up your bucket as fast as possible. Ok?"

At this point, the experimenter initiated the experimental trials. In all three conditions, 30 rounds of the game were played. In the Regular and Risk conditions, participants placed their cards face-down on the table, and the experimenter ensured they were not made visible until both players had made a choice. Both cards were then simultaneously turned over, upon which the experimenter called out the cards played, and counted out the appropriate points for each player into their respective buckets (e.g. "Two foxes. That's three points for (player 1) and two points for (player 2)"). In the Sequential condition, participants took turns going first, meaning that both had 15 turns as both first and second player to move. At the experimenter's prompting, cards were played face-up so that the second player to move could take the first player's choice into account when making his or her decision (e.g. "Ok, (player 1), you first this time. That's a fox. Your turn, (player 2). That's a lion. That's one point for (player 1) and one point for player 2").

6.2.5: Data analysis

On the first round only, participants' decisions were considered independent, since they could not have been influenced by knowledge of their opponents' strategies. Consequently, on this round only, the usual demographic predictors were used in the analysis (age in months, gender, relationship), in addition to experimental condition.

For analysis of decision-making over multiple rounds, the dyad rather than the individual was used as the unit of analysis, since individual participants' decisions were not fully independent of those of their partner. When analysing data over multiple rounds, the predictors included were dyads' mean age in months, gender (dummy-coded: all-male/all-female/mixed), relationship (known/unknown). In order to allow direct statistical comparisons between children at different stages of development, each dyad's mean age was used to categorize it into one of three age groups: 6-7 years (n=49), 8-9 years (n=54) and 10+ years (n=49).

6.3: Results

6.3.1: First Round Data

Out of 304 participants, 79% chose defect (i.e. their own card). Experimental condition did not affect this decision, and nor did any of the demographic factors, as shown by their failure to provide any significant predictive value in a stepwise binary logistic regression (opponent's card=1/own card=0). In a block entry logistic regression containing the same predictor and outcomes variables, all predictors failed to approach significance (all p-values >.4). This means that regardless of condition, age, gender, or the relationship with their partner, the majority of participants started the experiment with a competitive rather than a cooperative decision.

6.3.2: Multiple-Round Analysis

Over the course of 30 rounds dyads selected, on average, 11.35 Matches (SD=5.56), 17.17 Clashes (SD=6.33), and 1.41 Opposites (SD=1.8). The lowest number of Matches achieved was zero (chosen by three dyads), the highest was 28 (chosen by one dyad), and the modal outcome was 10 matches per dyad. Preliminary analysis showed that dyads' gender and relationship composition had no significant effect on Matches, Crashes or Opposites. Consequently, these variables are excluded from the proceeding analysis.

A factorial ANOVA including condition and age category as between-subjects factors showed that neither had an impact upon the total number of Matches over 30 rounds. Neither the main effects of condition ($F(2, 150)=2.53$, $p=.08$, ($\eta p^2=.034$) or age category ($F(2, 150)=2.06$, $p=.13$, $\eta p^2=.028$) achieved significance, nor was there a significant interaction between the two ($F(4, 148)=1.13$, $p=.34$, $\eta p^2=.031$). A second factorial ANOVA using the same between-subjects factors (condition and age category) similarly found that they also had no effect on dyads' overall number of Clashes. There was no main effect of condition ($F(2, 150)=.38$, $p=.68$, $\eta p^2=.005$) or age category ($F(2, 150)=.16$, $p=.86$, $\eta p^2=.002$), and no significant interaction between the two ($F(4, 148)=.73$, $p=.57$, P $\eta p^2=.02$).

Given the failure of the experimental manipulation of conditions in these initial analyses, the analysis of Opposites took on an important dimension as a means of task comprehension (see the Design section for a full explanation of this). The Opposites data was heavily positively skewed and included many scores of zero, meaning that consequently an ANOVA violated the assumptions of homogeneity of variance (Levene's test, $F(8, 143)=3.89$, $p<.001$). As a result, the Kruskal-Wallis test was instead used for this analysis, with Mann-Whitney analysis used to check post-hoc comparisons. The difference in opposites between the three conditions was significant (Kruskal-Wallis, $\chi^2(2)=20.09$, $p<.0001$). Follow-up comparisons between the three conditions showed that this difference was indeed due to lower levels of Opposites in the Sequential condition ($M=.58$, $SD=1.07$) compared to both the Regular ($M=1.59$, $SD=1.96$, $U=891.5$, $p=.004$) and Risk conditions ($M=2.04$, $SD=1.92$; $U=653.5$, $p<.00001$). This suggests that participants did not struggle to understand the scoring system, since they consistently avoided the lowest scoring outcome in the Sequential condition more than they did so in the other conditions. There was no significant difference between the number of opposites chosen in the Regular and Risk conditions ($U=1079$, $p=.13$).

The difference in the number of times dyads settled on either Matches or Clashes was then examined. Since there were no main effects of condition or age category generally, the entire dataset (152 dyads) was pooled for this analysis. A comparison of the total number of times that each dyad selected Match ($M=11.35$, $SD=5.56$) or Clash ($M=17.17$, $SD=6.33$) showed that the latter outcome was significantly more common ($t(150)=7.92$, $p<.000001$).

The next area of focus was how task iteration (as measured by trial number, from 1-30) affected the frequency of successful Matches achieved on each of the thirty trials. Within the whole dataset (i.e. all three conditions pooled, n=152), there was a moderate negative correlation between trial number and number of Matches, meaning that participants

converged on this higher-scoring outcome less frequently over the course of iteration (Spearman, two-tailed, $r_s(28) = -.52$, p=.002).

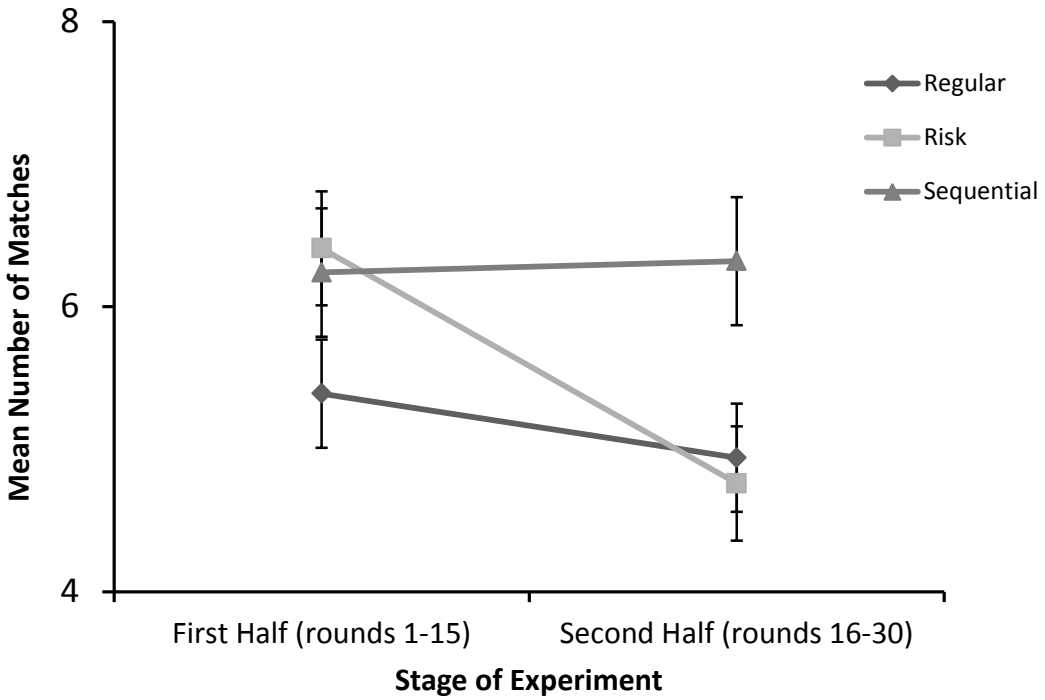

Figure 6: the interaction between condition and experimental stage.

In order to test this decline in successful cooperations directly, rounds were grouped into two blocks of 15 trials. These comprised of the first half stage (rounds 1-15) or second half stage (rounds 16-30) of the experiment, and a separate Match score for each stage was tallied for each dyad. A mixed-design ANOVA was performed with number of Matches in each experimental stage as a within-subjects factor and experimental condition and age category as between-subjects factors. Mauchley's test showed that the assumption of sphericity was violated in this data (p<.001), so consequently the Greenhouse-Geisser correction has been applied to all of the following results. There was a significant main effect of Stage (F(1)=8.76, p=.004, ηp^2=.058), a marginally significant main effect of condition (F(2, 143)=2.53, p=.083, ηp^2=.034) and a significant interaction between stage and condition (F(2)=5.71, p=.004, ηp^2=.074). There was no significant main effect of age category (F(2, 143)=2.06, p=.13, ηp^2=.028), and no two-way interactions between either age category and condition (F(4, 143)=1.14, p=.34, ηp^2=.031) or age category and stage (F(2)=.21, p=.81, ηp^2=.003), neither was there a three way interaction between age category, condition and stage (F(4)=.23, p=.92, ηp^2=.007). Age category thus had no effect on the number of Matches. The interaction between condition and experimental stage is shown in Figure 6.

As Figure 6 suggests, both the main effect of stage and the interaction between stage and condition were largely driven by a decline in Matches in the Risk condition from the first half to the second half of the experiment. There was no significant difference between Match levels in the three conditions in the first half stage of the experiment ($F(2, 149)=2.05$, $p=.13$). However, the difference had attained significance by the second half stage of the experiment ($F(2, 149)=3.22$, $p=.043$). Post hoc analysis revealed this result to be the product of a significantly higher frequency of Matches in the Sequential condition compared to those seen in Regular (LSD test, $p=.041$) and Risk (LSD test, $p=.022$). However, there was no significant change from first half to second half in the number of Matches in the Regular ($t(50)=1.1$, $p=.28$) or Sequential ($t(49)=-.21$, $p=.84$) conditions. By far the biggest change was observed in the Risk condition, in which the number of Matches declined sharply from the first stage to the second ($t(50)=5.42$, $p<.001$).

6.4: Discussion

Four hypotheses were tested in this experiment. **Hypothesis one** predicted an age effect whereby older dyads would achieve more successful cooperation than younger dyads. This was not supported. **Hypothesis two** predicted that children would adjust their cooperation levels according to the amount of information they had regarding their opponent's moves. This was also not supported, as overall cooperation in the Sequential condition was no higher than in either of the other conditions, although there was an interaction effect whereby this outcome had occurred by the second half of the experiment. **Hypothesis three** predicted that cooperation levels would be lowest in the Risk condition, due to increased fear of null payoffs. This too was not supported, since overall cooperation levels in the Risk condition were no different to the other two conditions. There was, however, an interaction between the Risk condition and task iteration whereby cooperation declined more steeply in this condition than in the Regular and Sequential conditions. **Hypothesis four** predicted that iterated play would lead to an increase in older children's cooperation levels over time, and have no effect on younger children's, which would remain consistently low. This was only partially supported. Younger children's cooperation levels did remain consistently low, but this pattern was also true of the cooperation levels of children of all ages.

In summary, the results of the BOS suggest that hyper-competitive motivations are sufficient to stifle cooperative behaviour throughout pre-adolescence, and that this is true even when

cooperative behaviour accords with the SEM. A reluctance to accrue a relative disadvantage to attain an absolute gain seems to be at the heart of this. This pattern was largely immune to experimental manipulation, or the prospect of strategic long-term cooperation presented by iteration. There was, however, evidence of iterated play allowing children learning to avoid zero scores in the Risk condition.

A degree of non-cooperation due to inequity aversion was anticipated, at least amongst the younger children tested. Nevertheless, the consistently low levels of cooperation across all conditions and demographic groups occurred to a greater degree than was expected. Mutual defections ("Crashes") significantly outnumbered Matches to a very high degree across all conditions, thus showing that hyper-competitiveness in children can violate the SEM by stymying cooperation in mutualistic, strategic settings. The frequency of Matches observed was far from negligible, but Matches were nevertheless heavily outnumbered by Crashes. This is a key finding.

A particularly surprising result was the absence of a main effect by which the Sequential generated higher levels of cooperation. This was the condition in which the second mover had perfect information as to the outcome of their decision on both players' payoffs, in contrast to the uncertainty in the other conditions. Whilst anticipating that sequential play would not eliminate hyper-competitiveness, I had expected a greater number of Matches in this condition, which at least afforded those subjects with positive-sum preferences the opportunity to easily act upon them, when compared to the Regular and Risk conditions. Instead, there was no difference between the overall frequencies of Matches across any of the three conditions, suggesting that an absolute gain which conferred a relative disadvantage was simply an unpopular preference, and that this fact was robust enough to resist experimental manipulation. It should be noted, however, that by the second half stage of the experiment, Matches were significantly higher in frequency in the Sequential condition than in the Risk and Regular conditions. This suggests that the removal of uncertainty from the procedure at least caused a lesser decline in the level of successful cooperation witnessed over time.

The predictions regarding the impact of Risk upon children's behaviour were also largely unsupported, despite the interaction shown in Figure 6 (p.132). The predicted age effects failed to appear, with all of the pre-adolescent dyads tested appearing to respond to risk in a broadly similar way. Whilst the prediction that cooperation would be significantly lower in the Risk condition overall was also not entirely borne out, there was clear evidence of an effect of iteration upon this condition, and it appears that by the later rounds subjects were beginning

to become significantly more averse to the dangers of costly miscoordination, as evidenced by the fact that they were making significantly fewer (risky) attempts at cooperation, a pattern which did not occur in the two conditions where miscoordinating cooperation was less of a hazard. The predicted pattern whereby Matches would be highest in the Sequential condition, intermediate in the Regular condition and lowest in the Risk condition did not materialise, although the interaction pattern visible in Figure 6 suggests that it may eventually have done so had the experiment been played over a greater number of rounds.

The effects of iteration upon cooperation levels were also less pronounced than expected, particularly in terms of the lack of interaction with dyads' mean ages. Instead of the predicted age effects, all dyads appear to have simply preferred mutual defection to compromise and cooperation at a fairly constant rate. Thus, children at the elder end of the age spectrum appear to have been motivated by DIA just as much as their younger counterparts. This means all age groups violated the SEM to an equally high degree. Whilst the lingering possibility of an unforeseen 0-0 payoff appeared to make subjects more averse to risk in the Risk condition, this learning affect appeared to be the only obvious effect of iteration. In the other two conditions, where a surprise payoff of 0-0 was not an issue, most subjects appear to have simply maintained a fairly consistent preference for not submitting to the possibility of relative disadvantage.

Previous findings on children's cooperative preferences in game theatrical contexts have found them to be adept cooperators in mutualistic, stag hunt-like contexts (Duguid et al., 2014; Grueneisen et al., 2015; Wyman et al., 2013), and fairly consistent defectors in more zero-sum scenarios based on the prisoner's dilemma (Angerer et al., 2015; Fan, 2000; Gummerum et al., 2009; Lergetporer et al., 2014; Sally & Hill, 2006). Both of these tendencies comply with the predictions of the SEM, since each strategy can be used to maximise income in its respective context. Surrounded by defectors in a PD and cooperators in an SH, one should match each strategy tit-for-tat. The BOS attempts to extend such findings by presenting subjects with a scenario in which a failure to cooperate is not merely competitive (as in the PD), but hyper-competitive, in that high levels of mutual defection harm the actor's overall payoffs as well as the recipient's. On the whole, subjects appeared to prefer this outcome to one of higher absolute but lower relative payoffs, thus providing strong evidence of a hyper-competitive preference interfering with potentially beneficial mutualistic cooperation.

This is an important finding because it demonstrates that just because children *can* cooperate remarkably well in a wide variety of contexts, this is does not mean that it can be assumed that

they *will* do so in all circumstances, even including those in which they should do according to the logic of self-interest. Whilst this experiment is the first to demonstrate such effects in a dynamic, strategic context, its findings are certainly not without precedent in the literature. It is becoming increasingly clear that a preference for rejecting disadvantageous inequity is extremely strong in childhood, and many children are more than willing to pay a cost to avoid it (Bereby-Meyer & Fiks, 2013; Blake & McAuliffe, 2011; Blake et al., 2014; Bügelmayer & Spiess, 2014; McAuliffe et al., 2014; Sheskin, Bloom, et al., 2014). Consideration of why this might be will be included in the general discussion (Chapter 8).

As in Chapter 5, there were no observed age differences in cooperation in this experiment. Previous developmental studies using the PD have been inconsistent in this area, with some finding an increase in cooperation in older children (Angerer et al., 2015; Fan, 2000; Gummerum et al., 2009) and others not doing so (Lergetporer et al., 2014; Sally & Hill, 2006). Previous work on the SH, meanwhile, has examined only pre-schoolers and therefore provides little clue as to what might be expected in terms of the development of mutualistic preferences (Duguid et al., 2014; Grueneisen et al., 2015; Wyman et al., 2013), something which the next chapter will attempt to rectify. The present study does not contain any instances of cohorts becoming actively *less* cooperative (and thus more likely to violate the SEM) in tandem with age, and in this sense at least it is not radically out of kilter with previous findings.

Perhaps slightly more surprising is the lack of an age effect in relation to children's risk sensitivity. Although the development of risk aversion in childhood is a complex process, the only recent comprehensive review of this literature concluded that, all other things being equal, risk aversion increases steadily from the pre-school years to early adulthood (Boyer, 2006). Nevertheless, it appeared that, in the context of the BOS specifically, dyads of all ages were able to learn to play more conservatively in the presence of risky outcomes at a fairly similar rate. It is, nevertheless, interesting that the presence of risk did not dampen dyads' cooperativeness as much as might be expected; instead, the frequency of Matches seemed reasonably impervious to experimental manipulation.

One thing which it is very important to highlight is that a lack of task comprehension is very unlikely to be the chief cause of the present experiment's results. Crucially, in the Opposites condition where the Clash outcome was both uniquely undesirable *and* easy to avoid, it was avoided to a much greater degree than when it was only uniquely undesirable (Risk) or neither of these things (Regular). Indeed, in the Sequential condition the Opposites outcome occurred

on fewer than 2% of trials, whereas if participants had been making random choices it would have occurred on 25%. Additionally, the Clashes scenario, which had equal outcomes which scored higher in absolute terms than Opposites, occurred on 56% of trials in the Sequential condition, thus showing that participants clearly preferred an outcome which payed higher in absolute terms (i.e. 1-1 instead of 0-0), so long as it did not mean their partner gaining an advantage in relative terms.

One potential criticism of this experiment which does bear scrutiny, however, is the suggestion that it may have had an implicitly competitive framing, and that this may have contributed to the relatively low cooperation levels seen in the results. As the Procedure section of this chapter makes clear, terms such as “win”, “player” and “game” were scrupulously avoided when creating the experiment’s script. Nevertheless, it could be argued with some justification that any scenario involving “cards”, “points” and “prizes” has an intuitively game-like feel, and may therefore have activated competitive schemas in many participants. This is probably a valid criticism and perhaps goes some way to explaining the very large disparities between the findings of the present experiment and previous examinations of children’s proclivities for mutualistic cooperation (Duguid et al., 2014; Grueneisen et al., 2015; Wyman et al., 2013). Nevertheless, it should be stressed that strong emphasis was placed on the experiment’s positive-sum potential throughout the BOS script, and it was clearly demonstrated that Matches filled the prize buckets the fastest, both factors which should have helped to mitigate competitive demand characteristics.

6.4.1: Conclusion

At the beginning of this chapter I expressed a desire to extend the conception of “cooperation” beyond simple resource allocation, and instead look at how cooperative and hyper-competitive preferences held up in a dynamic, strategic context. As the first ever developmental experiment to do this, the present study is only a first, tentative step towards exploring this topic, but as such it adds a number of novel findings to the literature and opens up many new avenues for future research. On the whole, the results are clear. Although non-negligible, the levels of cooperation in the BOS were low when one considers that adherence to the SEM requires them to be extremely high in this experiment. Particularly important in demonstrating this are the results from the Sequential condition, in which the low number of Matches achieved can only be attributable to a deliberate refusal to accept relative disadvantage in the interests of attaining an absolute material gain. This finding goes beyond

merely demonstrating the existence and strength of hyper-competitiveness, as in previous chapters, instead giving a concrete example of how it can dominate even profitable, mutualistic cooperation in a developmental sample. Humans are indeed uniquely proficient cooperators and these abilities undoubtedly begin to emerge early in the lifespan (Tomasello, 2014). However, the evidence that we are also unusually fierce competitors in our early years continues to grow too.

Chapter 7: The Stag Hunt and Chicken Games

7.1: Introduction

In the previous chapter, children across an age spectrum ranging from pre-school to late childhood largely failed to establish long-term cooperative outcomes in a mutualistic cooperation task. The motivations behind this appear to have been hyper-competitive. This chapter therefore seeks to examine whether such motivations extend even further and undermine mutualistic cooperation in the stag hunt. The stag hunt is the most basic of all of the 2x2 cooperation dilemmas, both in terms of how easy it is for cooperation to emerge (Archetti & Scheuring, 2012), and the cognitive demands it places upon its players (Mueller, 2003). More so than almost any other economic experiment, then, the stag hunt is not one in which hyper-competitive motivations should be expected to stifle cooperation. Previous developmental studies with stag hunts, or at least stag hunt-like parameters, have shown that, indeed, pre-school-aged children can often solve this dilemma and achieve mutual cooperation (Duguid et al., 2014; Grueneisen et al., 2015; Tomasello et al., 2012a; Wyman et al., 2013). However, there is no experimental data on any children other than pre-schoolers, and also no data on how iterated play affects children's cooperation. This chapter is therefore an exploratory attempt to broaden our knowledge of children' mutualistic cooperation by extending both the age range of those tested, and the complexity of the scenarios they are exposed to, in order to see what makes mutualism thrive and what causes it to fail. Furthermore, data will also be collected from the chicken game, an "anti-coordination" game which is essentially a reverse stag hunt, in order to provide extra insight into children's stag hunt preferences.

7.1.1: Aims and Objectives

1) To further examine children's hyper-competitive preferences in a mutualistic cooperation task, this time using a simpler experiment than that used in the previous chapter, namely the stag hunt.
2) To explore differences in how children solve a mutualistic cooperation task under conditions of full information and limited information as to the task's payoff structure.

3) To explore how children's mutualistic cooperation develops over repeated interactions in an iterated stag hunt.
4) To apply the stag hunt to a wider age spectrum than has previously been attempted. There is no existing data on how people between the ages of pre-school and adulthood approach this experiment.
5) To use the chicken game, an anti-coordination game, to shed light on some of the motivations underlying children's behaviour in the stag hunt.

7.1.2: The Stag Hunt

The stag hunt (SH) is a simple metaphor with far-reaching implications for human cooperation and pro-sociality. In recent years, anthropologists and philosophers alike have suggested that the types of dilemma it models may lie at the root of our species' unique cooperative abilities (Skyrms, 2004; Tomasello et al., 2012a). It has also been suggested that our ability to cooperate in the types of mutualistic scenario represented by the SH metaphor is innate (McLoone & Smead, 2014). The stag hunt analogy is as follows: two hunters are out searching for meat. Working alone, both have the potential to catch themselves a hare, a prize of modest value. If, and only if, they work together, they will have the collective ability to bring down a stag, a prize of much greater value. However, if only one hunter pursues the stag, and the other works alone to pursue a hare, the former hunter will be left with nothing, the worst possible outcome. Thus, it makes sense to pursue the big prize only if one is confident of the acquiescence of one's fellow hunter. For an example of how this translates into game theoretical payoffs, see Game A in Table 17.

Table 17: payoff matrices for (A) the SH, and (B) the CG.

(A) Stag Hunt

		P2	
		Cooperate	Defect
P1	Cooperate	3, 3	0, 1
	Defect	1, 0	1, 1

(B) Chicken Game

		P2	
		Cooperate	Defect
P1	Cooperate	2, 2	1, 3
	Defect	3, 1	1, 1

As the previous chapter explained, the cooperation represented by the SH is not altruistic, but mutualistic (Baumard et al., 2013; Tomasello et al., 2012a). It is therefore cooperative but not hyper-cooperative; it does not violate the SEM. No altruism or reciprocity is required

(Tomasello, 2014), and thus no violation of the SEM is required. This makes satisfying the conditions for cooperation in the stag hunt far easier than in other game theoretic paradigms such as the prisoner's dilemma, as well as making cooperation in the SH a less cognitively demanding task than in other 2x2 games (Archetti & Scheuring, 2012; Mueller, 2003).

On the whole, mutualistic outcomes were limited in the BOS experiment (Chapter 6), but the SH is an easier experiment for cooperation to emerge in. Unlike the BOS, no compromise is needed. Both parties score equally well when they successfully achieve joint cooperation, so there should be no reason for inequity aversion to interfere with cooperative intent. The SH therefore represents an interesting ultimate test of the limits of hyper-competitive ORPs, since it is the game theoretical paradigm in which they should be in the shortest supply. A SH player following the SEM has few if any reasons to compete (defect) as long as s/he suspects that the same is true of her partner. The primate literature provides an instructive example here. Chimpanzees and bonobos have consistently proved themselves to be rational maximisers across a wide range of adapted game theoretical and behavioural economics paradigms, since they fail to deviate from the SEM by showing any other-regarding preferences, either positive or negative (Jensen et al., 2007a; Jensen et al., 2006; Kaiser et al., 2012; Riedl et al., 2012). Presented with the SH, however, chimpanzees can prove to be adept cooperators (Bullinger et al., 2011; Duguid et al., 2014). This finding shows how, in practice as well as in theory, positive ORPs are not necessary for successful cooperation in this experiment.

Although not quite as selfish as higher primates, pre-school children are also not characterised by high levels of other-regarding behaviour. This statement is true of both altruistic acts (Fehr et al., 2008; House et al., 2013) and punitive acts (Bereby-Meyer & Fiks, 2013; Takagishi et al, 2014; Wittig et al, 2013; see also Chapters 3 and 4). Interestingly, however, they too have shown a strong proclivity to reach cooperative outcomes in the few previous SH experiments performed to date (Duguid et al., 2014; Wyman et al., 2013), and even outperformed chimpanzees when the two species were directly tested against one another using comparable methodologies (Duguid et al., 2014). In both of these studies, pre-school children's success was aided considerably by the use of communication. However, they have since also been observed successfully solving other conceptually similar mutualistic coordination puzzles in the absence of any communication as to their partner's intent (Grueneisen et al., 2015).

There is an important caveat to the consistently mutualistic findings of this growing literature, however, illustrated nicely by a recent study by Brosnan and colleagues (2011). This experiment, which tested the successful cooperative abilities of adult humans alongside both

higher primates and monkeys in the SH, showed than mutualistic cooperation dopes not simply emerge naturally in all contexts. Specifically, they showed how it struggled to occur under conditions of uncertainty. The researchers, in order to make their methodology as comparable as possible across multiple species, implemented a procedure whereby subjects of all species were not informed as to the experiment's payoffs, and were simply left to deduce these payoffs for themselves through trial and error over repeated rounds.

All non-human primate species largely failed to "solve" the SH and settle on stag-stag (i.e. mutual cooperation) under such circumstances. Importantly, however, only 19% of adult human dyads succeeded in converging upon the stag-stag payoff in this experiment. This shows that, whilst it may be a superficially simple task, solving the SH in the most mutualistic manner is far from trivial in the absence of full information and external guidance. Indeed, even when presented with clear instructions, many adults still fail to converge on the mutualistic outcome in the SH. Experimental studies on the SH are few, but one recent study found that even when informed as to the payoffs more than 40% of adult dyads failed to consistently converge upon the stag-stag outcome over the course of repeated rounds, with many dyads instead adopting the safer but less profitable hare-hare pattern. Having honed in on either strategy, dyads' behaviour tends to stay there quite consistently, meaning that both patterns become more frequent with iteration, as the stag-hare mismatch scenario tends to die out (Al-Ubaydli, Jones, & Weel, 2013).

7.1.3: The Present Chapter

The work to date on the SH represents an intriguing starting point, but a starting point nonetheless. Compared to the prisoner's dilemma, probably the most studied paradigm in all of game theory, there is very little experimental SH data. This seems strange, since the mutualistic scenario presented by the SH is just as plausible as the harshly zero-sum one embodied by the PD. The present study aims to use a series of SH experiments to test whether mutualistic cooperation is robust in relation to the hyper-competitive preferences which children are prone to.

This chapter consists of four experiments, comprising three stag hunts (Experiments 5a, 5b, and 5c) and one chicken game (Experiment 5d). In all of the SH experiments, the payoff matrix used was that shown as Game A in Table 17 (p.140). Experiment 5d was based on the chicken game, shown as Game B in the same table. Experiment 5a pairs children with a helpful

confederate in the form of a parent, and looks at how children behave in the SH under conditions of full and absent information as to the experiment's payoffs. Experiment 5b examines how all-child dyads perform in a scenario with full information. Experiment 5c compares how children perform with two types of helpful adult confederate, a parent and a stranger. Experiment 5d explores all-child dyads' preferences in a chicken game with full information, in order to compare them to the preferences shown in the SH.

Such areas remain largely unknown, so these experiments are exploratory and my predictions tentative. Nevertheless, the high levels of cooperation exhibited by the pre-school participants in previous studies (Duguid et al., 2014; Wyman et al., 2012) are not expected to be exhibited across the whole range of all conditions and experiments employed here. Whether or not any age effects will emerge remains unclear. Linear age effects have been observed in some studies of middle childhood (e.g. Blake & McAuliffe, 2011), but not others (e.g. House et al, 2013). Generally, however, when such effects do occur, they tend to show that older children are more prosocial than younger peers (Eisenberg et al., 1998), so this will form the basis of several working hypotheses in the absence of anything to suggest otherwise.

How children's preferences for cooperation respond to repeated interactions will also be interesting. Unlike many classic economic experiments such as the public goods game (Chaudhuri, 2011; Ledyard, 1995), overall cooperation does not necessarily decline with task repetition in the stag hunt, although participants tend to converge on repeatedly hunting either stag/stag or hare/hare (Al-Ubaydli et al., 2013; McLoone & Smead, 2014). Children are known to be capable of withdrawing cooperation conditionally (Angerer et al., 2015; Lergetporer et al., 2014), and can do so in iterated tasks from as young as 5-years old (Vogelsang et al., 2014), so it seems possible that they will exhibit strategic changes in cooperation over the course of repeated interactions.

7.2.1: Experiment 5a – the Stag Hunt with Explained and Unexplained Payoffs

In Experiment 5a, each dyad consisted of a child aged between 4-8 years and their parent. Dyads were tested in one of two conditions. These were the Unexplained condition, in which dyads had to work out the SH payoffs for themselves, and the Explained condition, in which the payoffs were explained to them before the experiment began. This design aimed to incorporate aspects from two of the key previous studies outlined in the introduction (Brosnan

et al., 2011; Wyman et al., 2013), in order to maximise the comparability of this experiment's findings to those of the existing literature. These two aspects were the use of adult confederates (i.e. the parents), and the two experimental conditions.

The adult confederate aspect was adapted from Wyman et al, 2013, in which 4-year old subjects were paired with an adult confederate who, in the experimental condition, was instructed to try and help the child by signalling the intent to cooperate on certain trials on which the stag-stag outcome was attainable. In the present experiment, although they were not encouraged to explicitly signal their intentions, parents were essentially used to fulfil this role of helpful confederate. Whilst there were no material rewards for parents, there were for the children, and it was assumed that parents would be intrinsically motivated to help their young children succeed as much as possible.

The Unexplained payoffs condition is a novelty borrowed from Brosnan et al, 2011, who used it in a comparative SH study of adults, apes and monkeys to demonstrate that success at the stag hunt is non-trivial even for humans when the parameters of the task have to be uncovered through reinforcement learning. Thus, although always paired with a helpful confederate (the parent), children were nevertheless presented with conditions whereby a cooperatively-minded individual should find settling on the cooperative equilibrium (stag-stag) to be either achievable (Explained) or difficult (Unexplained). This was done in order to compare how able and willing children were in each condition to initiate and persist with cooperation rather than competition. It will be a matter of considerable interest to see how they fare when uncertainty and guesswork are involved (i.e. in the Unexplained condition). This has the potential to cause children to experience unfavourable outcomes, thus calling into question whether cooperation could emerge and survive amongst such potential setbacks.

7.2.1.1: Hypotheses:

Experiment 5a was designed to test the following hypotheses:

- **(H1)** Successful mutualistic cooperation will be significantly higher in the Explained than in the Unexplained condition.
- **(H2)** No age effects are expected in the Explained condition, in which cooperation should be consistently high, as in previous developmental studies with transparent payoffs.

- **(H3)** In the Unexplained condition, it is expected that dyads featuring older children will be more likely to converge upon the stag-stag solution, due to the children's superior strategic thinking abilities.
- **(H4)** No effects of iteration are expected in the Explained condition, in which cooperation should be consistently high.
- **(H5)** Iteration is expected to cause increasing cooperation in the Unexplained condition, as dyads figure out the highest paying outcomes over time.

7.2.2: Methods

7.2.2.1: Participants

Subjects were 118 children aged 4-8 years (M=6.28 years, SD=1.75; 59 boys/59 girls), each paired with a parent (n=118, 55 fathers/63 mothers). All participants were recruited at the Live Science Area in the same manner as in the other experiments in this dissertation (see Chapter 2). Table 18 shows the demographic details of subjects, both sorted by condition and combined into a single sample.

Table 18: descriptive demographic statistics for Experiment 5a.

	age in years		gender		parent	
	M	SD	Male	Female	Father	Mother
Overall	6.28	1.75	59	59	55	63
Unexplained	6.35	1.84	28	30	35	23
Explained	6.21	1.33	29	31	20	40

7.2.2.2: Materials

The materials and experimental set-up were very similar to the BOS game described in the previous chapter (see Figure 5 in Chapter 6, Section 6.2.2, p.126 for a visual representation). In both conditions, children and parents were issued with the same materials. Each was given two decks of 22 playing cards, one featuring cartoon images of lions and the other featuring foxes. Within each dyad, one of these animals represented cooperation and the other defection (order counterbalanced). Participants in each dyad were sat facing opposite one another at a table. Each had an identical, sturdy, wooden occluder placed in front of them, in

order to obscure their un-played cards and corresponding choices from their partner. Payoffs from each round were represented by counting plastic gumball capsules ("points") into transparent plastic containers adjacent to each participant. These were exchanged for prizes at the end of the experiment in the same manner as in previous chapters (see Chapter 2) for details.

7.2.2.3: Design

Experimental Conditions and Payoffs

The design was between-subjects, with dyads assigned to one of two conditions, "Explained" or "Unexplained". Both conditions involved the same SH payoffs, shown below in Table 19. The key difference is that in the Explained condition participants were informed as to these payoffs before the experimental trials began, whereas in the Unexplained condition they received no such information. Full details of the differences between the habituation methods used in each condition are explained below in the Procedure section. In both conditions, 20 iterations of the SH were played. This lesser number of rounds compared to the previous BOS game (Chapter 6) was employed because the present study involved younger children, who it was felt might struggle to maintain their focus over a longer experiment.

Table 19: the stag hunt payoff matrix used in Experiments 5a, 5b, and 5c.

		P2	
		Cooperate	Defect
P1	Cooperate	3, 3	0, 1
	Defect	1, 0	1, 1

Outcomes, Dependent Variables and Scoring

The most beneficial outcome for both participants in all dyads was when they both chose cooperate, an outcome which is the key dependent variable and will henceforth be referred to as Stag. A second important dependent variable was when both subjects chose to defect, a payoff which will henceforth be referred to as Hare. If one subject within the dyad cooperated and the other defected, then payoffs were not equal, as the defecting player would score one and the cooperating player would score zero. When the defector in this scenario was the child,

this outcome is referred to as Child Advantage. When the defector in the scenario was the parent this scenario is referred to as Parent Advantage. All of these outcomes are illustrated by Table 19.

7.2.2.4: Procedure

After volunteering, children and parents were sat opposite one another and issued with their cooperate and defect cards. These were easily accessible to each participant, but well hidden from their partner by their occluder. The transparent buckets were placed on the table. It was explained to participants that they were about to take part in an experiment, the aim of which was to fill the bucket with as many points as possible. At this stage, the prizes were shown to participants, and the number of points necessary to win each prize was explained. Just as in the previous BOS experiment (Chapter 6), it was heavily stressed to participants that it was their absolute and *not* their relative score that would win them better prizes[8].

There then followed an explanation of scoring. In the Unexplained condition, dyads were told that points were scored by playing different combinations of cards, but that the experimenter could not tell them exactly how many – instead, they would have to work it out for themselves. Participants were simply told to play one card face-down each time the experimenter said the word "go", and then turn it over when the experimenter said "turn". They were encouraged to watch carefully how many points each combination of animals scored, as demonstrated by the number the experimenter counted out into each participant's bucket in each case.

In the Explained condition, dyads were asked to play each different paired combination of animal cards in turn (order counterbalanced). In each case, the experimenter then physically counted out the appropriate number of points into each participant's bucket. All payoffs to both partners were counted out in this manner slowly, clearly and publically, so that both were in no doubt as to the payoff to their partner as well as to themselves. The four possible outcomes (i.e. Stag, Hare, Child Advantage, and Parent Advantage) were demonstrated twice to all dyads in this condition. After this demonstration, child subjects were asked to explain back to the experimenter how many points each combination of cards scored. In cases where they could not do this, the demonstration trials were repeated until they could. No dyads were

[8] "It does not matter how many points Mum/Dad gets, ok? Mum/Dad's score does not affect your prize."

allowed to begin the experimental stage until children could clearly and without hesitation state the payoffs for each combination of cards back to the experimenter.

Before beginning the 20 experimental rounds, the experimenter emptied participants' prize buckets of any points accumulated during the habituation trials. Participants were told that they were not practicing any more, and that it was now time to "do it for real". At this point, it was *again* heavily stressed to participants that better prizes were accumulated by scoring in absolute rather than relative terms[9].

Twenty rounds of the stag hunt experiment were then played. All rounds involved cards being played in the same manner as in the habituation trials, with both players initially placing a card face down in the middle of the table, and then simultaneously revealing their choices to one another at the experimenter's prompt ("turn"). At the end of 20 trials, points were tallied, and children were allowed to choose their prize accordingly. Participants were then thanked and debriefed.

7.2.3: Results

7.2.3.1: First Round Analysis

At a *dyadic* level, there was no significant difference between the observed and expected count of successful Stag outcomes between the two conditions. In the Explained condition the observed count of Stags on the first round was 19 (31.7%), and in the Unexplained condition it was also 19 (32.8%), $\chi^2(1)=.016$, p=.899.

At an *individual* level (i.e. focusing on children's cooperate/defect choices), however, a stepwise binary logistic regression of child's first round behaviour (outcome: cooperate=1, defect=0) using three predictors (age in months; child's gender; experimental condition) found only an effect of condition, whereby children were more likely to display a *lesser* level of cooperation in the Explained condition (β=-.89, Wald=5.62, p=.018). This was in direct contrast to predictions, as it means greater information regarding payoffs made children more likely to

[9]"Remember that very important thing I told you earlier: Mum/Dad's score *does not affect your prize*. To get a big prize, you just need to fill up your bucket as fast as possible. Ok?"

defect. This is reflected in the descriptive statistics. In the Unexplained condition there were 35 cooperators (60%), whereas in the Explained condition there were just 23 (38%).

7.2.3.2: Repeated Rounds Analysis

Preliminary analysis found that none of the demographic variables predicted overall cooperation levels. A stepwise linear regression analysis featuring total stag coordinations as the outcome variable, revealed no significant predictive effect of child's age in months, child's gender or parent's gender. A block entry regression containing all of these factors did not approach significance (R^2=.013, F(3, 114)=.49, p=.69; p-values for all predictors >.3).

Contrary to predictions, there was no significant difference in the mean number of achieved stag coordinations between the two conditions. In the Explained condition, M=5.47, SD=4.41; in the Unexplained condition, M=4.88, SD=4.12 (t(116)=-.743, p=.46). This appears to have been because children's levels of cooperation were, on the whole, not high. When data from the sample as a whole was pooled, children cooperated an average of just 7.42 times (SD=4.53) out of 20.

As in Chapter 6, the experiment was divided into two stages, first half and second half, to see if cooperation levels changed in response to task repetition. However, the above pattern remained consistent throughout the experiment. A mixed ANOVA was performed with condition and age category as between-subjects factors and experimental stage as a within-subjects factor. There was no main effect of condition (F(1, 108)=1.6, p=.21, ηp^2=.015), no main effect of age category (F(4, 108)=1.05, p=.38, ηp^2=.037) and no main effect of stage (F(1, 108)=.055, p=.82 ηp^2=.001). There were also no significant two-way interactions between age category and stage (F(4, 108)=1.03, p=.4, ηp^2=.037), or condition and stage (F(1, 108)=.089, p=.77, ηp^2=.001), and no significant three-way interaction between the three variables (F(4, 108=.99, p=.41, ηp^2=.036).

Analysis then shifted to explore differences in occurring frequencies between the four dyadic outcomes explained in the design section: Stag, Hare, Child Advantage, and Parent Advantage. The frequencies of each of these payoff outcomes (from 0-20 in all cases, due to there being 20 experimental trials) were compared within-subjects for each dyad (see Figure 7). The data for parental advantage was heavily positively skewed. Consequently, non-parametric measures were used for this analysis. Friedman's ANOVA showed a significant difference between the

mean frequencies of the four outcomes, $\chi^2(2)=90.81$, p<.000001. Post-hoc, pairwise Wilcoxon comparisons revealed this to be due to significantly higher numbers of Child Advantages (M=7.13, SD=3.12) than there were Stags (M=5.18, SD=4.28, Z=-3.9, p<.0001), (M=.5.45, SD=3.94, Z=-3.67, p<.001) or Parental Advantages (M=2.25, SD=2.14, Z=-8.13, p<.000001). There were also fewer Parental Advantages than Stags (Z=-5.83, p<.000001) and Hares (Z=-6.41, p<.000001). There was no significant difference between the latter two, however (Z=-.399, p=.69).

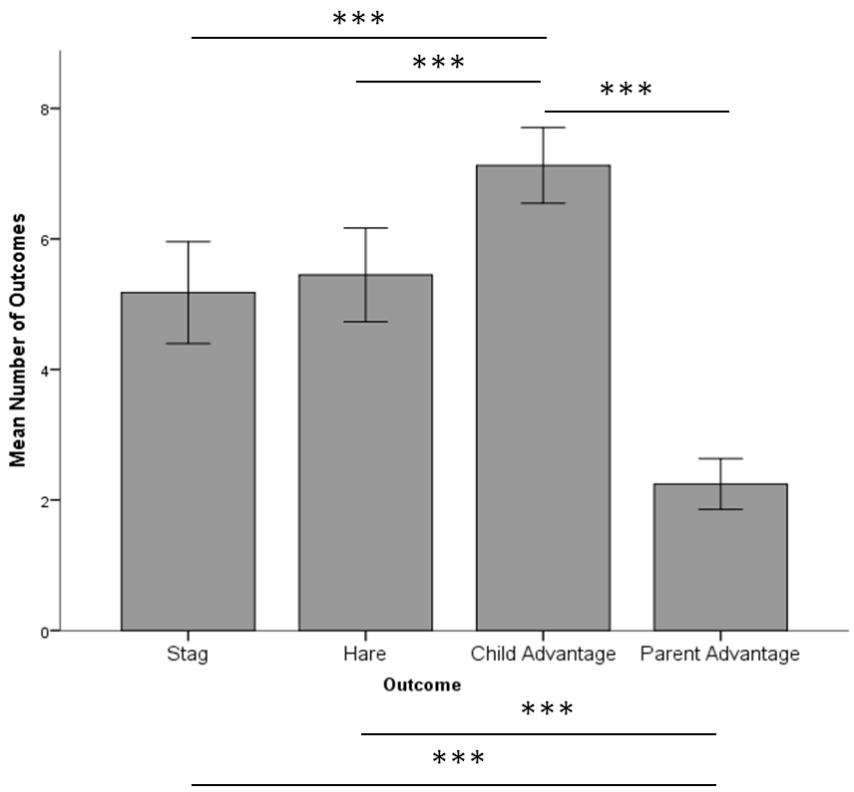

Figure 7: mean number of times out of 20 that outcomes in E1 occurred.
Error bars show 95% confidence intervals.

As Figure 7 shows, these data clearly suggest that children were preferentially choosing defection over cooperation, since the frequency of outcomes resulting from child defection (Hare and Child Advantage) exceeds the frequency of outcomes resulting from child cooperation (Stag and Parental Advantage). The total number of times each child chose to defect was calculated by adding each dyad's total number of Hares and Child Advantages. Children's overall defection levels did indeed dominate their overall cooperation levels, and children defected on an average of 12.58 rounds (SD=4.53), a figure significantly higher than chance (i.e. 10 rounds; t(117)=6.18, p<.000001). This finding is equally solid within both conditions when they are analysed discretely (within Explained, M=12.88, SD=4.64, t(59)=4.81, p<.0001; within Unexplained, M=12.26, SD=4.43, t(57)=3.89, p<.001).

7.2.3.3: Adult Behaviour in Experiment 5a

Analysis of adult preferences was performed in order to check that adults would, as predicted, behave in a way which suggested that they were motivated to help their children score highly. This was done by looking at whether they cooperated more in the Explained condition, in which the way to help their children was more obvious from the outset, than in the Unexplained condition, where they did not initially have this knowledge to guide them.

The overall mean number of cooperative decisions out of 20 by adults was 12.31 (SD=4.36). Adults did indeed make more cooperative decisions (over 20 rounds) in the Explained (M=13.18, SD=3.96) as opposed to the Unexplained condition (M=11.4, SD=4.58, t(116)=-2.26, p=.026). Both of these figures are significantly greater than chance, i.e. 10 cooperative choices (Explained, t=(59)=6.22, p<.001; Unexplained, t(57)=2.32, p=0.24).

Adults were more cooperative from the outset in the Explained condition. On round 1 they made 47 cooperative choices (81%), whereas in the Unexplained condition they made only 27 (47%), a figure which was significantly lower ($\chi^2(1)=12.74$, p<.001).

Stepwise linear regression analysis was performed with the cumulative number of cooperative decisions made by children as the outcome variable. The predictors were the total number of cooperative decisions made by parents (preliminary analysis showed that there was no significant difference between the behaviours of mothers and fathers), in addition to the three demographic variables of age, gender and relationship. Since there was no main effect of condition, the data from all subjects was pooled in this regression. There was found to be one significant model (overall model, R^2=.235, F(1, 116)=35.58, p<.001). The only significant predictor within this model was the total number of adult cooperations, which had a positive effect (β=.504, t=5.97, p<.001). This means that, to some extent at least, children were adjusting their strategy rationally, since they were more likely to try for the Stag payoff when their partner was doing the same.

7.2.4: Mini-Discussion Experiment 5a

Five hypotheses were tested in Experiment 5a. **Hypothesis 1**, that successful mutualistic cooperation would be significantly higher in the Explained condition than in the Unexplained condition, was not supported. **Hypothesis 2** was only partially supported. Although there were no age effects in the Explained condition, this was due to consistently low cooperation as opposed to consistently high cooperation. There was no support for **Hypothesis 3**, as there were no age effects in the Unexplained condition, in which cooperation was also generally not high. **Hypothesis 4** was only partially supported; although there was no effect of iteration upon cooperation in the explained condition, this was due to cooperation being consistently low as opposed to consistently high. **Hypothesis 5** was not supported. Cooperation did not increase over time in the Unexplained condition.

As the failure or partial failure of these hypotheses demonstrates, children were unexpectedly uncooperative in this stag hunt experiment. Competitiveness rather than mutualistic cooperation was exhibited by children throughout the sample. Virtually none of the pre-experiment predictions were supported. Crucially, the experimental manipulation was ineffective, and there was no overall difference in cooperation between the two conditions. The exception to this was the first round, on which full payoff comprehension (Explained) made children *less* cooperative, in direct contrast to what was predicted. This means that directly having had the stag hunt's mutualistic payoffs explained to them, children responded by defecting significantly more than those who were playing blindly. Cumulatively, the results of Experiment 5a showed strong evidence of hyper-competitiveness in the child sample, as competition dominated cooperation even though subjects were partnered with a broadly helpful confederate. Outcomes whereby the child hunted the hare whilst the adult hunted stag were the most common. The lack of age or gender effects seems to suggest that this overall pattern was true to a similar extent across the entire sample, which ranged throughout early to middle childhood, from 4-8 years of age.

The lack of a difference in overall success-rates in achieving the stag/stag payoff between the two conditions initially represents something of a surprise, but may be explicable in terms of how the two levels of payoff comprehension differently affected children and adults. Whilst adults were significantly more cooperative from the start when informed of the payoff structure, the higher levels of initial adult cooperation were offset by lower levels of child cooperation, which explains the overall similarity in the frequency of Stag payoffs.

Another interesting finding is the results of the regression showing that adults' overall cooperation frequencies significantly predicted children's overall cooperation frequencies. This

does suggest both an ability and a degree of willingness on the part of children to rationally adjust their level of cooperativeness in response to that of their partner's, a finding suggestive of a degree of contingent cooperation (House et al, 2013).

Amongst adults, although cooperation dominated competition, the mean cooperation level (12 out of 20) is arguably not the completely overwhelming majority that might be expected if adults' only motivation had been trying to maximize their children's payoffs. This finding instead suggests that other motives may also have present amongst parents. Indeed, when quizzed about this afterwards many adults answered that they were keen to occasionally "test" their child or "keep them guessing", in order to keep the procedure interesting for them. This slightly random element to their behaviour is not typical of what one would expect from a participant fully motivated only by the task payoffs, and may have contributed to children's relatively low cooperation levels, by violating trust and encouraging contingently non-cooperative responses to occasionally unhelpful behaviour. It was therefore very possibly an inhibitory factor in allowing many dyads in the sample to move towards a stable equilibrium. Experiment 5b thus sought to rectify this by performing the stag hunt with all-child dyads, in which neither subject was a confederate.

7.3: Experiment 5b – the Stag Hunt with all-child dyads

The results of Experiment 5a clearly appeared to suggest that children in early to middle childhood favoured defection in the stag hunt. This is an intriguing yet unexpected finding, and therefore requires further investigation. Children in Experiment 5a were, effectively, paired with largely helpful but slightly unpredictable stooges, and thus the experiment was more psychological than game theoretical. Experiment 5b sought to test children's real economic preferences in the stag hunt, unencumbered by the potential confound of being paired with a partner unconcerned with their own material payoffs. To this end, all participants in Experiment 5b were children, with dyads being as closely matched for age as possible. Given the failure of the Explained and Unexplained conditions to produce radically different cooperative scenarios in the previous experiment, all dyads in this experiment had payoffs fully explained to them, in order to examine their cooperative tendencies under conditions of as much comprehension as possible. In a scenario in which both participants had a genuine material interest in accruing points, would cooperation be more likely to occur? Previous experimental work in a one-shot stag hunt-style experiment would appear to suggest so

(Duguid et al., 2014). However, the results of Experiment 5a appear to suggest that cooperation in Experiment 5b would be relatively low.

7.3.1.1: Hypotheses

Hypotheses in Experiment 5b were based upon the findings in Experiment 5a, and therefore adopted the assumption that many children would prefer to defect in the SH. Cooperation was expected to be intermediate, due to some children targeting high scores through mutualism, and others defaulting to competitive behaviour.

- **(H1)** As a result, Experiment 5b will witness the classic bimodal pattern of cooperation seen in a typical stag hunt with adults. Dyads will converge upon a stable equilibrium of Stag or Hare, with the intermediate patterns dying out over the course of repeated rounds.
- **(H2)** Should any age effects occur, these will take the form of greater cooperation, and therefore more Stag outcomes, in older children.

7.3.2: Methods

7.3.2.1: Participants

Subjects were 122 children (therefore 61 dyads) aged 4-8 years. Participants were recruited in the usual manner. A full demographic overview of the dyads participating in Experiment 5b is shown in Table 20.

Table 20: descriptive demographic statistics for dyads in Experiment 5b.

mean dyad age (years)		gender		relationship		
M	SD	Male	Female	Mixed	Known	Unknown
5.99	1.06	18	20	23	33	28

7.3.2.2: Design and Materials

All dyads sat the same stag hunt experiment, effectively the Explained condition from the previous experiment, with the same payoff structure and habituation system. The only difference was that prizes were now given to both participants at the end of the experiment, rather than having one participant (i.e. the parent) acting as a stooge. All materials were identical to those in Experiment 5a. The key outcome variables were the same as in the previous experiment, namely the overall frequencies of Stags and Hares, and how these changed over time. Children's choices to cooperate or defect on the first rounds (i.e. binary, yes/no), were also of interest.

7.3.2.3: Procedure

The procedure was basically analogous to that of the Explained condition in Experiment 5a. Partners were positioned opposite one other and materials were distributed in the same way. The differing payoff permutations resulting from each combination of cards were systematically explained to all participants during habituation. These were repeated as often as necessary for both children in each dyad to be able to fluently repeat these payoffs back to the experimenter prior to the commencement of the experimental trials. The absolute rather than relative nature of the scoring system was stressed at the same two stages in the habituation as in the BOS and in SH Experiment 5a. Twenty experimental rounds were again played.

7.3.3: Results

7.3.3.1: First Round Analysis

Looking first at individual first round decisions, only 41 participants (34%) chose to cooperate. As in Experiment 5a, individual cooperate/defect choices on round 1 were analysed, in order to examine participants' baseline behaviour prior to their experiencing of their opponents' strategies. Stepwise binary logistic regression found no significant predictors of this outcome. A block entry binary regression containing the same predictors found that none of them approached significance (all p-values >.15). This means that the entire sample, regardless of age, gender or relationship, preferred to defect on round 1 to an equivalent extent.

7.3.3.2: Repeated Rounds Analysis

Preliminary stepwise regression analysis was used to examine whether any of the demographic factors predicted Stag outcomes within dyads. There was one successful model (R^2=.85, F(1, 59)=5.5., p=.022), which contained relationship. Dyads that were aquatinted with one another prior to the experiment ("known") were significantly more likely to coordinate on Stag (β=1.62, t(59)=2.35, p=.022). None of the other demographic factors provide significant (all p-values >.3).

The total number of Stags versus the total number of Hares over 20 rounds was compared within-dyads. Hares (M=12.61, SD=4.72) hugely outnumbered Stags (M=1.92, SD=2.78; t(60)=-11.82, p<.00001), meaning that defection dominated cooperation in the overall experiment.

The data were split into two stages (first half/second half) in order to examine whether this overwhelming tendency to coordinate hare instead of stag varied throughout the experiment. There was no significant change in the number of Stags (out of 10) from first (M=1.08, SD=1.56) to second half (M=.9, SD=1.67); t(60)=.85, p=.39) of the experiment. Interestingly, however, the number of Hares increased significantly as the experiment progressed from first (M=5.57, SD=2.28) to second half (M=7.03, SD=2.92; t(60)=-5.01, p<.001). This means that only one of the predicted equilibria was increasingly settled upon, i.e. the Hare equilibrium.

This in turn would suggest that the frequency of cooperative decisions from players was declining as the rounds went on. Indeed, a look at the frequency of individual subjects' cooperative decisions on round 1 shows that 41 participants (33.6%) chose to cooperate. This percentage was significantly below chance (binomial, test prop=.5, observed prop=.34, p<.001), yet it still represents something of a high water mark for cooperation in Experiment 5b as a whole when compared to what follows. A Spearman correlation (2-tailed), reveals a strong, negative relationship between increasing round number and the frequency of total individual decisions to cooperate, r_s(18)=.788, p<.0001 (see Figure 8).

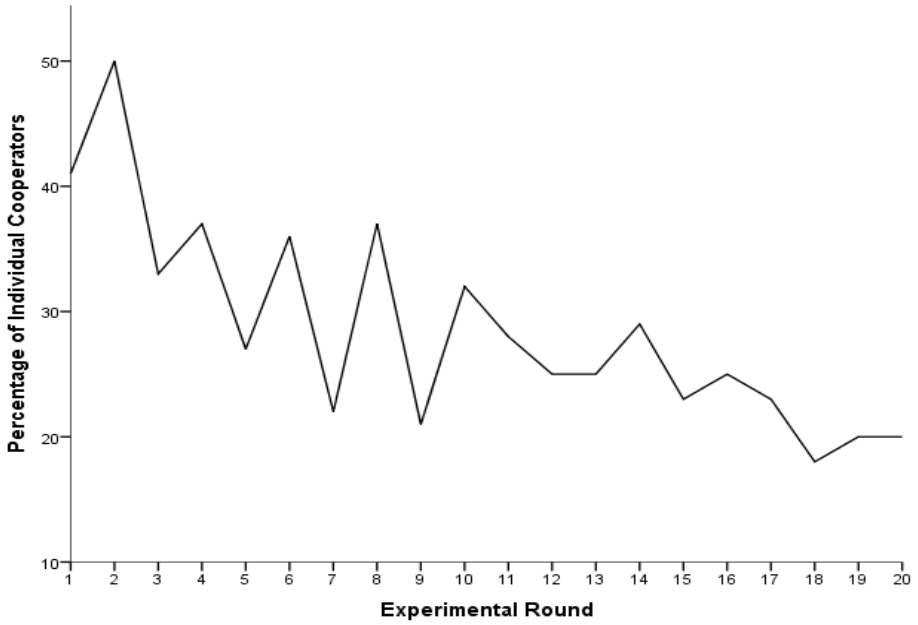

Figure 8: round-by-Round Percentage of Individual Cooperators in Experiment 5b.

Out of 61 dyads, the highest number of Stags scored out of 20 was 10, and this was achieved by a solitary, outlying dyad. Clearly, then, the predicted division of the sample into cooperative (Stag) and defecting (Hare) dyads also failed to occur. Instead, dyads overwhelmingly converged on the latter (73% had arrived at Hare by round 20). Hare outcomes increased in frequency, but a convergent increase in Stags did not. Indeed, fully 29 out of 60 dyads (48%) matched nothing but Hare from round 15 onwards, and Hare was the majority outcome over the final five rounds for 53 dyads (88%). Once again, there were no age effects, either in terms of individual first round choices or overall dyadic play. Although there was, as detailed above in the regression analysis, a relationship effect whereby known dyads were more likely to attain Stag outcomes than strange dyads, it should be noted that even amongst known dyads these frequencies were not high. Even in these dyads, Stags totalled only 14% of round outcomes (the overall mean number of Stags for the sample as a whole was 10%).

7.3.4: Mini-Discussion Experiment 5b

Two hypotheses were tested in Experiment 5b. **Hypothesis 1** was largely unsupported. There was no bimodal pattern of cooperative outcomes. Although the predicted greater frequency of Hare equilibria with iteration did occur, this was due to children massively preferring defection to cooperation, meaning that the corresponding increase in the Stag equilibrium failed to occur. **Hypothesis 2** was also not supported. As in Experiment 5a, there was no effect whereby

older dyads were able to achieve more cooperative Stag outcomes. There was, however, an effect of relationship, whereby dyads who were known to one another (friends and siblings) achieved higher overall frequencies of Stag outcomes than dyads who were strangers.

The results of Experiment 5b unequivocally show competition dominating cooperation, with the Hare outcome outstripping the Stag outcome by a ratio of 6:1 over just 20 rounds. This time this outcome occurred in a scenario with real incentives for all players. Once again, the extent to which defection dominated cooperation was unexpected, and meant that the experimental hypotheses had to be largely rejected. Far from being "intermediate", cooperation levels amongst children were once again below chance levels even from the first round, and they continued to decline thereafter. The predicted subset of dyads who managed to converge on the Stag equilibrium failed to appear.

Not all of the subjects in this experiment adopted defection as their default strategy, as the results of round 1 show. Instead, 34% of subjects attempted to achieve the Stag outcome by cooperating on round 1. This is clearly a minority, but nevertheless a substantial one. Amongst this minority, from around round 3 onwards, there was clear evidence of conditional cooperation. This is best illustrated by Figure 8, which shows a gradual, conditional pattern of decline in cooperative overtures amongst potential cooperators when they are exposed to defections over repeated rounds. It closely resembles the types of patterns of cooperation which are frequently witnessed in classic public goods games (Chaudhuri, 2011; Fehr & Gächter, 2002; Fischbacher et al., 2001; Ledyard, 1997). Until recently it was not known for certain whether young children were capable of adopting this strategy, but the present study is one of the clearest contributions yet to a growing literature showing that this behaviour is present from very early in childhood (Alencar, Siqueira, & Yamamoto, 2008; Angerer et al., 2015; Lergetporer et al., 2014; Vogelsang et al., 2014). It should be noted that, although cooperating in the SH is potentially the best outcome, withdrawing one's cooperation when faced with a defecting partner is an entirely sensible thing to do, and adheres to the SEM. Less obvious is the motivation underlying the behaviour of the majority of subjects who chose to adopt defection from the outset. Experiment 5d will attempt to probe this motivation by introducing subjects to the chicken game.

Before that, however, one final result of interest in this experiment requires exploration. Regression analysis showed that being pre-acquainted with one's partner appeared to slightly mitigate the tendency towards defection. Although known dyads' (friends and siblings) cooperative choices were still dominated by defection, this tendency occurred significantly less

than amongst dyads of strangers, whose levels of trust in one another appear to have been particularly low.

7.4: Comparing Experiments 5a and 5b

Participants in Experiment 5b appear to have been even less cooperative than in Experiment 5a. Despite the differences in the types of samples used, a number of comparisons between the results of Experiments 5a and 5b were attempted, in order to quantify the apparent differences in the ways which children played the SH against adults and peers. One potential confound to these comparisons identified in the analysis of Experiment 5b is that known participants achieved the cooperative Stag coordination more than strangers. This confound will be explored in more depth in Experiment 5c. However, since cooperation amongst even the known dyads in Experiment 5b was low, some comparison between the two experiments might still be instructive.

The total number of Stags was compared between the child/adult dyads in the Explained condition in Experiment 5a, and the all-child dyads in Experiment 5b. In other words, the cases in which both partners knew the payoffs were compared across experiments. Comparing the frequency of stag coordinations achieved by dyads in the Explained condition in Experiment 5a (n=60, M=5.47, SD=4.41) with those achieved by the dyads in Experiment 5b (n=61, M=1.98, SD=2.78) confirmed that, as expected, the former was significantly larger. Levene's test showed a violation of the assumption of homogeneity of variance between the data from the two experiments ($F(1, 119)=7.57$, $p=.007$), so the non-parametric Mann-Whitney test was employed to demonstrate this ($U=927$, $p<.001$).

The pattern of lesser cooperation in Experiment 5b is present from the first round. Chi-square analysis showed that stags were significantly more likely to be achieved on round one by adult/child (observed count=32%) over child/child dyads (observed count=15%; expected count=25%; χ^2 (1)=9.18, $p=.002$). In addition, binomial analysis (1=stag, 0=other, test proportion=.25) showed that child/adult dyads were likely to achieve stag coordinations on round 1 at significantly above-chance levels (i.e. above .25; observed prop=.32, $p=.047$), whereas child/child dyads achieved stag coordinations at significantly *below*-chance levels (observed prop=.11, $p=.007$). This suggests than children default to a less cooperative strategy when interacting with peers than with adults.

In summary, being paired with a peer appears to make children even less likely to cooperate mutualistically than being paired with an adult.

7.5: Experiment 5c: the Stag Hunt with strange and familiar adults

Experiment 5c was designed to compare the extent to which children cooperated in the SH when paired with their parents and with strange adults. This was done for two reasons. The first reason is that Experiment 5b exposed a potential confound in the results of Experiment 5a. This was that known dyads achieved the cooperative Stag coordination more than the strange dyads. This renders direct comparison with Experiment 5a slightly problematic, since all dyads in that experiment (parent/child) were obviously well-known to one another, whereas in Experiment 5b approximately half of the dyads were strangers. This may therefore explain at least as much of the variance as is explained by the use of all-child dyads. In other words, the reason for the lower cooperation in Experiment 5b may be that children do not trust strangers, rather than that they do not like cooperating mutualistically with other children. The second motivation underlying Experiment 5c was some of the slightly unusual behaviour of the parents in Experiment 5a. Of particular note in this respect was the fact that many of them seemed to deliberately alternate their cooperate and defect choices in order to "test" their children. This, naturally, would have affected children's responses.

It is hard to say exactly how this arrangement might have prejudiced the experimental outcome, precisely because it is an unusual one within developmental psychology experiments, although at least one previous study suggests that children are no less generous towards their own parents in economic games than they are towards strangers (Peters, Ünür, Clark, & Schulze, 2004). On one hand, the results of Experiment 5b suggest that, if anything, being paired with a familiar partner ought to have inflated children's levels of cooperation. On the other, being paired with a smiling parent in a paradigm with "points" and "cards" may have made the children more prone to interpreting the scenario as a game, and activating an appropriately competitive schema.

In order to investigate these possibilities, Experiment 5c consisted of a second experiment with child/adult dyads, this time attempting to investigate whether being paired with their parents caused children to behave in an idiosyncratic manner, either in terms of inflating or suppressing cooperation. The working assumption adopted was that (H1) children would

behave more cooperatively towards their parents than they would towards a strange adult in the SH.

7.5.1: Methods

7.5.1.1: Participants

Participants were 36 children aged 4-11 years (mean age = 6.85 years, SD = 2.03, 21 boys/15 girls). Also recruited were 18 parents (8 males/10 females). Demographic statistics are given in Table 21.

Table 21: descriptive statistics for child participants in Experiment 5c.

	age in years		gender	
	M	SD	Male	Female
Overall	6.85	2.03	21	15
Parent	6.85	2.01	10	8
Confederate	6.85	2.12	11	7

7.5.1.2: Design and Materials

Children were assigned randomly and equally to one of two conditions, between-subjects. In the Parent condition, they were paired with a parent, as in Experiment 5a. In the Confederate condition, they were paired with a female confederate. In both conditions, all payoffs were fully explained to all participants.

The same setting for the experiment was used as in Experiments 5a and 5b. For children, the same card decks were used as in these experiments. Unlike these previous experiments, however, the adult in Experiment 5c was given a single, rigged deck, with 12 cooperate and 8 defect cards shuffled in a random order. These frequencies were chosen as 12 cooperations was the mean number chosen by adults in Experiment 5a (see Section 7.2.3.2, p.149). Standardising adult cooperation numbers at this level thus allowed any observed differences in cooperation levels between conditions to be attributable to partner-type (Parent/Confederate) as opposed to partner's cooperativeness, since children demonstrated an ability in both

Experiment's 5a and 5b to adapt their cooperation levels contingently, depending upon the cooperativeness of their partner.

7.5.1.3: Procedure

Most aspects of the procedure were identical to the previous experiments in this chapter, including the materials, instructions to children, habituation, number of rounds (20) and prize structure. The only important departure was in the use of the rigged adult card decks. This feature of the experiment was kept secret from the child participants, who were implicitly led to believe that their opponent's preferences were genuine (see below). Children therefore thought they were playing against an opponent making real choices, as had been genuinely the case in Experiments 5a and 5b.

Instead of the adults being given two adjacent piles of cards, they were given a single pre-shuffled deck behind their occluder, on top of a sign saying "please play these cards in the order given." As the parent sat down, the experimenter tapped this sign firmly with their finger (this gesture was hidden behind the occluder from the child's perspective) in order to make sure the parent's attention was drawn to it. All parents managed to successfully follow this instruction without moving their deck from behind the occluder or asking any comprehension questions which might have revealed the deception to their child. Confederates were fully briefed as to the nature of the deception, and thus simply played their cards in the order given.

Children were given cooperate and defect cards split into two decks and given a free choice between each on every round, just as in previous experiments in this chapter and Chapter 6. They were led to believe that their adult partner had been presented with the same scenario, and that their adult partner's preferences were their own, as opposed to pre-ordained by the experimenter as they were in reality.

7.5.2: Results

For Parent dyads, the mean number of stag coordinations achieved was 3.06 (SD=2.34), whereas for Confederate dyads this figure was 4.5 (SD=3.05). Partner type (Parent/Confederate) did not affect the frequency of Stag coordinations ($t(34)=1.59$, $p=.15$).

This pattern was consistent throughout the experiment. A mixed ANOVA with stage as a within-subjects factor and partner-type as a between-subjects factor showed no main effect of partner type ($F(1, 34)=2.54$, $p=.12$, $\eta p^2=.069$), no main effect of stage ($F(1, 34)=.491$, $p=.49$ $\eta p^2=.014$) and no interaction between the two ($F(1, 34)=.87$, $p=.36$, $\eta p^2=.025$).

7.5.3: Mini-Discussion Experiment 5c

Hypothesis 1 was not supported. Children's cooperation levels when paired with adults did not vary between parents and strangers when the adult's cooperation levels were kept constant.

Whilst acknowledging the modest sample size, there is little or no suggestion from this experiment that children cooperate radically more or less with parents than with strangers in this task. There is no indication in these data that being paired with a familiar adult would have grossly confounded the Experiment 5a data away from an alternative result that would have occurred had they been paired with a confederate or other stranger. The lack of a conditional decline in cooperation, as well as the general pattern of relatively few overall Stag coordinations (19% here, compared with 26% in the first experiment), were both replicated and did not differ by condition. Collectively, the data from the first three experiments suggest that most children prefer to defect in the stag hunt, and that this is true in a wide range of conditions and amongst varied types of opponent. When attempting a mutualistic cooperation task, children seem particularly mistrustful and non-cooperative towards unfamiliar peers.

7.6: Experiment 5d – the Chicken Game

Having ruled out potential confounds in the findings thus far, the dominance of defection behaviour amongst children in the data now requires further investigation. To do this, Experiment 5d will co-opt a different game theoretical paradigm, namely the chicken game (CG). The payoffs of the CG and SH are presented for comparative purposes once again in Table 22.

The prisoner's dilemma is the most challenging of the simple 2x2 game theoretical paradigms for cooperation to emerge in, since two motivations must be overcome if one is to cooperate rather than defect. Firstly, there is one's own greed, and secondly there is the fear of one's partner's greed (Bornstein, Mingelgrin, & Rutte, 1996). In the SH, cooperation is less

challenging, since only fear should cause defection, whereas greed will actually promote cooperation, as it is now the latter that can potentially result in the best outcome. However, this is only the case if one assumes that the actor evaluating the dilemma is "rational" in a purely economic sense. One potential explanation of the data in Experiments 5a-5c (and indeed the BOS in Chapter 6) is that instead of prizing absolute payoffs, many children are unable to resist evaluating the SH in relative terms. In other words, they are defecting because they want to "win" the SH, even though doing so is irrational according to the SEM.

Table 22: Payoff matrices for (A) the SH, and (B) the CG.

(A) Stag Hunt

		P2	
		Cooperate	Defect
P1	Cooperate	3, 3	0, 1
	Defect	1, 0	1, 1

(B) Chicken Game

		P2	
		Cooperate	Defect
P1	Cooperate	2, 2	1, 3
	Defect	3, 1	1, 1

This potentially intriguing idea may be testable. If the actor rather prizes relative advantage over absolute payoffs, as I suspect many of the children in the previous three experiments did, then defecting and trying to "win" the stag hunt also makes sense for greedy reasons. In effect, then, many of the children in Experiments 5a-5c may have been treating the stag hunt as a prisoner's dilemma, and may thus have been motivated by both greed and fear.

To test this theory, Experiment 5d introduced the chicken game. In the CG, an "anti-coordination" game, the stag hunt scenario is reversed, i.e. it is greed which motivates defection, and fear which motivates cooperation. This mirroring of the SH allows the CG to be used as a useful evaluative tool of children's SH motivations. If the children in Experiments 5a-5c really were defecting in the SH due to a subjectively-perceived "greedy" desire to "win" the game, then we would expect to see very high levels of defection in the CG. If, on the other hand, they were playing conservatively out of fear in order to avoid the null-payoff, their CG behaviour should be characterised by very high levels of cooperation. If they are motivated by neither of these cognitions, then the SEM predicts that a rational player with no knowledge or inferences as to their opponent's intentions should cooperate and defect to an equal extent in the chicken game, since neither strategy is automatically preferable.

7.6.1: Hypotheses

It was these three possibilities which Experiment 5d set out to investigate. Based on previous findings stressing children's predilection for relative advantage (Sheskin, Bloom, et al., 2014), and relatively high risk tolerance when compared to adults (Boyer, 2006), I tentatively attribute their SH preferences thus far primarily to greed, and thus predict that mutual defection will predominate over mutual cooperation in the CG. Therefore:

- **(H1)** Dyads in the Chicken Game will mutually defect ("Crash") significantly more than they will mutually cooperate ("Draw") in the CG.
- **(H2)** If any age effects are present, they will appear as a lesser preference for mutual defection pattern in older dyads.
- **(H3)** If any effects of relationship are present, they will appear as a greater preference for mutual defection amongst known dyads.
- **(H4)** Theoretically speaking, no equilibrium is likely in the CG, so neither cooperation nor defection will come to dominate the experiment in its later rounds.

7.6.2: Methods

7.6.2.1: Participants

Subjects were 104 children aged 4-12 years (M=7.55, SD=2.21; 48 boys, 56 girls). Experiment 5d had a within-subjects design, with all dyads participating in the same chicken game. Descriptive statistics are shown in Table 23. Considerations of design, materials and procedure were essentially identical to Experiment 5b. The only difference was, of course, that the habituation trials reflected chicken game payoffs rather than stag hunt payoffs.

Table 23: descriptive statistics for dyads in Experiment 5d.

mean age (years)		dyad gender			dyad relationship	
M	SD	Male	Female	Mixed	Known	Unknown
7.55	2.21	12	16	24	29	23

7.6.2.2: Design

There were a few new design features in this experiment which need to be explained. First, of course, is that the payoff matrix now reflected a chicken game rather than a stag hunt (i.e. Game B in Table 22, as opposed to Game A). Because of this, the outcome variables were also somewhat different. Those of interest are mutual cooperation and mutual defection. In the CG, mutual cooperation is termed a "Draw", and pays two points to both members of a dyad. Mutual defection is termed a "Crash", and is the worst possible outcome, paying both subjects zero points. When partners fail to coordinate decisions in either direction, the defecting partner benefits more from the Round ("Win"; 3 points), and the cooperator gets a low payoff ("Loss", 1 point).

For the purposes of the statistical analysis of age effects, dyads were grouped into three age categories of approximately even size according to their mean age: 5-6 year olds (n=17), 7-8 year olds (n=17), and 9+ year olds (n=18).

7.6.3: Results

7.6.3.1: First Round Analysis

At a *dyadic* level**,** binomial analysis revealed no difference between the expected (prop.=.25) and observed frequencies of either Draws (observed prop=.23, p=.45) or Crashes (observed prop.=.29, p=.31). At an *individual* level, stepwise logistic regression of subjects' decisions (cooperate=1/defect=0) on round 1 showed no predictive effects of any of the demographic variables. A block entry logistic regression confirmed that none of them approached significance (all p-values >.5). This suggests that the first round was very much in line with what classical game theory would predict. With no clues as to how their opponent would play, most subjects were essentially deciding on a strategy at random, and this was true across all demographic sub-sets.

7.6.3.2: Repeated Rounds Analysis

Over 20 rounds, dyads converged upon Crash (M=8.12; SD=3.56) more often than Draw (M=2.79, SD=2.86). Unsurprisingly, given the large disparity between these means, Crashes were observed to be the significantly more common outcome (Z=-5.16, p<.000001; Wilcoxon test used to adjust for violations of homogeneity of variance). This means that, contrary to

predictions, a dominant strategy emerged in the chicken game, and this strategy was defection.

Two stepwise regressions were performed to see if, firstly, total Draws per dyad, and, secondly, total Crashes per dyad, could be predicted demographically. In both cases, mean age in months was the only significant predictor. Increasing mean age predicted fewer draws (β=-.04, t=-2.59, p=.012; overall model, R^2=.119, F(1, 50)=6.73, p=.012), and a greater number of crashes (β=.064, t=3.93, p<.001; overall model, R^2=.236, F(1, 50)=15.42, p<.001). This means that contrary to predictions older children were more likely to defect than younger children in this experiment. The hypothesised effects of relationship with partner were not a factor in the chicken game.

As in previous experiments, the experiment was split into two stages reflecting the first half (rounds 1-10) and second half (rounds 11-20). The frequencies of each dyad's Draws and Crashes within each stage were then tallied. For Draws, a mixed ANOVA was performed with stage as a within-subjects factor and age category as a between-subjects factor. There was a significant main effect of age category (F(2, 49)=3.32, p=.044, ηp^2=.119), but no main effect of experimental stage (F(1, 50)=1.58, p=.22, ηp^2=.031). There was a marginally significant interaction between these two variables (F(2, 49)=2.62, p=.083, ηp^2=.097). The Greenhouse-Geisser correction was used upon all results in this ANOVA, as Mauchley's test indicated a violation of sphericity (<.001).

Follow up analyses were conducted to explore these findings further. There was only a marginally significant difference in the number of Draws between the three age categories within the first half stage of the experiment (F(2, 48)=.292, p=.064), which post-hoc testing showed to be down to 5-6 year olds achieving a significantly higher number of draws (M=2.12, SD=1.9) than those aged 7-8 years (M=1, SD=.89; LSD test, p=.022). Those aged 9+ (M=1.39, SD=1.04) did not differ significantly in their frequency of draws from those in the other two conditions (LSD tests, p>.1). Within the second half stage of the experiment, the difference in the frequency of draws between the three conditions was fully significant (F(2, 48)=3.42, p=.041). Post hoc testing showed this to be due to a significantly higher frequency of Draws (M=2.46, SD=.59) amongst those aged 5-6 as compared to those aged 9+ (M=.56, SD=.62; LSD test, p=.012). There was no significant difference between the frequency of Draws attained by those aged 7-8 years (M=1.31, SD=1.54) as compared to those in the other two age categories (LSD tests, p>.2). Comparing frequencies between the first and second halves of the experiment, Draws exhibited no change amongst the 5-6 years category (t(16)=.16, p=.88), or

the 7-8 years category (t(16)=-.51, p=.62). However, the frequency of Draws declined significantly amongst those in the 9+ category (t(17)=3.07. p=.007). Thus, the interaction appears to have been driven by the reduction in the number of Draws amongst those in the eldest age category, whereas the main effect of age seems to have been chiefly attributable to consistently higher frequencies of Draws amongst the youngest age group (5-6 year olds). The youngest age group was, therefore, consistently the most mutually cooperative in this experiment.

The same analyses were then applied to Crashes. A mixed ANOVA with stage as a within-subjects factor and age category as a between-subjects factor found a marginally significant main effect of stage (F(1, 50)=3.42, p=.065), a significant main effect of age category (F2, 49)=12.28, p<.0001) and a significant interaction between the two (F(2, 49)= 5.68, p=.006). The interaction is shown in Figure 9.

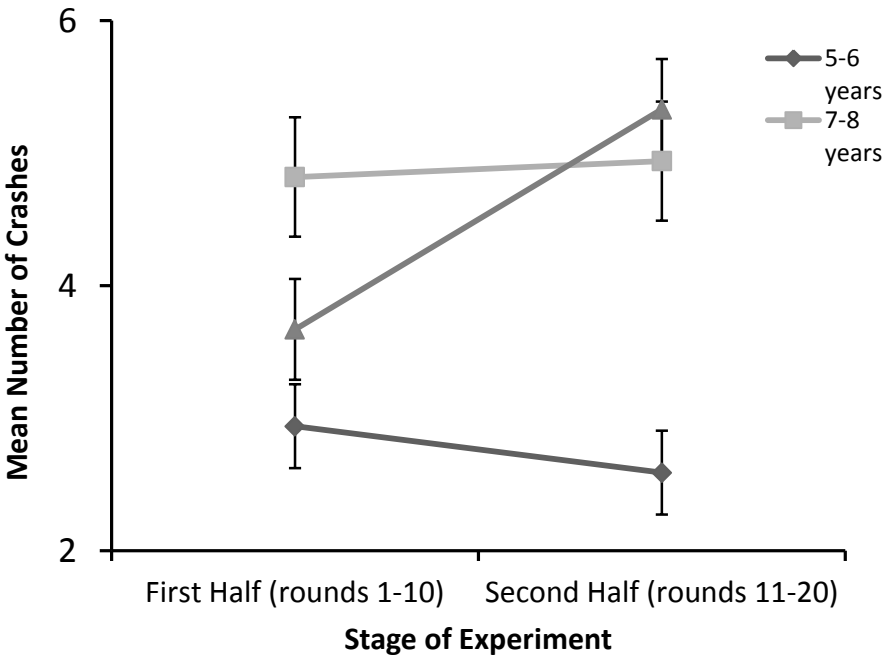

Figure 9: interaction between age category and experimental stage in chicken game

As the Figure shows, the interaction was driven primarily by a large increase in the number of Crashes amongst dyads in the 9+ years age category. Within the first half stage of the experiment, the difference in Crash levels between the age categories was significant (F(2, 49)=6.39, p=.003), and this was due to a higher frequency of crashes amongst dyads in the 7-8 years category (M=4.82, SD=1.74) than in either the younger 5-6 years category (M=2.94, SD=1.6; LSD test, p=.001) or the eldest 9+ years category (M=3.67, SD=1.23; LSD test, p=.032). There was no significant difference in Crash levels between the 5-6 years category and the 9+

years category (LSD test, p=.17). In the second half stage, the difference in the frequency of Crashes between the age categories was also significant (F(2, 49)=13.2, p<.0001), and post hoc testing showed that this difference was attributable to much lower frequencies of Crashes amongst the youngest 5-6 age cohort (M=2.59, SD=2.09) as compared to the intermediate 7-8 years cohort (M=4.94, SD=1.52; LSD test, p<.001), and also the 9+ years cohort (M=5.33, SD=1.41; LSD test, p<.0001). There was no significant difference between the mean scores of the elder two groups (LSD test, p=.49). Crash levels remained constant amongst those in the cohorts aged 5-6 years (t(16)=.59, p=.56) and 7-8 years (t(16)=-.31, p=.76), but increased significantly in those aged 9+ years (t(17)=-5, p<.001). These results for both Draws and Crashes ran counter to almost all predictions. Overall what this suggests is that children in late childhood exhibit the most hyper-cooperative tendencies in the chicken game, and that this tendency increases with task repetition, whereas younger children play the CG in a more consistently conservative manner.

7.6.4: Mini-Discussion Experiment 5d

Four hypotheses were tested in this experiment. **Hypothesis 1** was supported. Dyads in the game were much more likely to mutually defect than to mutually cooperate. **Hypothesis 2** was not supported. Contrary to predictions, it was the younger children who were more likely to cooperate in the CG, and the eldest age cohort appeared to become less inclined to cooperate as time went on. **Hypothesis 3** was not supported. There was no effect of relationship on the frequency of cooperative outcomes**. Hypothesis 4** was not supported. Contrary to the prediction that neither mutual cooperation nor mutual defection would become the dominant strategy, the latter appeared to be increasing in frequency as the experiment wore on, at least amongst the eldest age cohort.

The number of Crashes amongst dyads hugely outstripped the number of Draws, at a ratio of around 4:1. This means that subjects in most dyads prioritised adopting a risky strategy and trying to “win” the CG over a defensive strategy which sought merely to avoid crashing. However, this finding was accompanied by two other important but unexpected findings. Firstly, the age effect visible in Figure 9 is contrary to that predicted. Younger children were actually more cooperative than their elder counterparts in this experiment, an unusual result. Additionally, there was a very interesting interaction between age and experimental stage which, contrary to my prediction of no systematic fluctuations in the frequency of cooperation,

showed that the eldest age group in this experiment action became radically less cooperative with task iteration.

The dominance of defection in the CG seems telling. This is clearly hyper-competitive behaviour. Continuing to defect when your opponent does the same is totally counter-productive behaviour in the CG, and results in a very low score. According to the SEM, any "rational" actor should therefore avoid persisting with defection and switch to cooperation when this happens. On the whole, however, this did not happen, and the extremely low mean Draw frequency (14% of trials) suggests that when Draws did occur, either one or both partners within the dyad quickly abandoned the cooperative strategy in pursuit of another win. This behaviour unambiguously shows greedy motivations dominating fearful ones. Consequently, this leads me to speculate that the SH behaviours seen in Experiments 5a-5c are unlikely to be attributable to fear – at least not initially – and may well therefore be the consequence of subjects with a zero-sum attitude trying to "win" the SH, and prioritising this relative victory over mutualism.

The interaction between dyads' age category and experimental stage is also worthy of further comment. In the general introduction to this dissertation it was explained that although linear age effects are by no means a feature of all experiments on the development of cooperation, one fairly consistent finding is that, if and when they do occur, they almost always show elder children behaving more cooperatively. In the present experiment, however, there is an unusual example of the opposite effect. This is most likely explicable in terms of the SEM. The youngest children (the 5-6 years cohort) adhered most closely to the SEM in Experiment 5d, prioritising neither cooperation nor defection at the expense of each other, which is the most rational strategy in a game with no equilibrium. By contrast, the older children were, as is the case in many experiments, more willing to deviate from it, and more willing to risk potential costs to do so. A fuller discussion of this point, encompassing all of the wider findings in this dissertation, will be made in the General Discussion (see Chapter 8, Section 8.2.1, p.182). Whilst greater deviation from the SEM with age may explain the main effect of age, the interaction effect whereby the eldest cohort (9+ years) became more likely to defect in later rounds is still worthy of note. It appears that dyads in this cohort became increasingly bloody-minded as the experiment progressed, engaging in a form of conditional cooperation which, in the context of this particular experiment, bordered on mutually assured destruction. Why this might be the case is not immediately obvious, but an attempt to replicate it would be an interesting avenue of future research. Certainly, task repetition seemed to have a particularly

pronounced effect in reinforcing hyper-competitiveness in those in late childhood in this particular anti-coordination game.

7.7: Chapter 7 Overall Discussion

The aim of this chapter was to explore whether the hyper-competitive ORPs of children in early to middle childhood (4-8) would disrupt cooperation in a mutualistic scenario, the iterated stag hunt. The overall findings across four experiments showed that they do, and furthermore that the extent to which they do is greater than might have been expected based on previous literature.

Collectively the results of Experiments 5a-5d suggest an intriguing conclusion: far from cooperating mutualistically, many of my subjects in this chapter appear to have been trying to "win" the stag hunt. This is a contentious statement. Not only does such behaviour run completely counter to what conventional game theory would class as optimal behaviour (Archetti & Scheuring, 2012), but it also contradicts the predictions of prominent recent anthropological work which places the SH at the heart of human social behaviour, and stresses children's natural proclivity for mutualistic cooperation (Baumard et al., 2013; Tomasello, 2014; Tomasello et al., 2012a). Additionally, it directly conflicts with actual empirical findings in support of these ideas (Duguid et al., 2014; Grueneisen et al., 2015; Wyman et al., 2013). Of course, trying to "win" a SH is economically illogical and violates the SEM. It does, however, make sense to an individual who values their outcome relative to their neighbour more than they value it in an absolute sense.

7.7.1: Key Findings

Whilst the consistently and reliably low levels of mutualism and cooperation across four separate experiments is a striking result in and of itself, there are three specific findings which support the suggestion that many children treated this particular class of stag hunts as a zero-sum exercise. First, in Experiments 5a and 5b, explaining the SH payoffs made children actively *more* likely to defect on the first round. In other words, immediately after they had had the different outcomes explained to them, and before they had any chance to be influenced by their partner's strategy, children opted to defect at a higher level than those playing blindly (i.e. those in the Unexplained condition), or when compared to chance (in Experiment 5c).

Second, the presence in Experiments 5a and 5c of broadly helpful confederates who gave them multiple hints and opportunities to revert to cooperation failed to tempt the majority of subjects away from preferring defection. At least 75% of SH subjects in these two experiments experienced the Stag payoff at least once in an experimental trial, yet there was no upswing in cooperation in the later stages of either experiment to suggest that this caused subjects to switch towards it. Third, whilst one possible explanation for the previous finding is that children were (justifiably) fearful that their partner would defect and thus did the same only to protect themselves, such an interpretation is almost impossible to square with the results of the chicken game. In the CG, most children showed a very cavalier attitude to the possibility of getting zero on a given round, and instead set about repeatedly risking this outcome in pursuit of one which gave them a relative advantage, in spite of its consistent unprofitability within the majority of dyads (i.e. repeated Crash scores of 0-0).

Collectively, the SH and CG results in this Chapter are not dissimilar to those on children's preferences in the prisoner's dilemma (Fan, 2000; Lergetporer et al., 2014; Sally & Hill, 2006), though the present study extends such findings considerably by showing that the non-cooperative preferences shown in the PD extend to scenarios in which cooperation should be far less demanding to achieve. In game theoretical terms, the non-cooperative outcomes chosen by most children were extremely "inefficient", as there was very little selfish or utilitarian maximization, and a strong tendency to frame the social dilemmas in zero-sum terms even though this was, objectively, needless and destructive.

7.7.2: The Effect of Iteration - Conditional Cooperation

This chapter, along with the previous one, sought to explore how children's cooperation would persist over repeated interactions. This question is rarely posed in the developmental literature relative to the adult literature, and perhaps surprisingly so given that it is a central concern for the latter. Due to children's initial preference for non-cooperation in the stag hunt, mean cooperation declined over time and the Stag equilibrium was largely neglected. Both of these findings are in contrast to what might be expected in an adult sample (Al-Ubaydli et al., 2013), and indeed to a sample composed of higher primates (Bullinger et al., 2011; Duguid et al., 2014).This notwithstanding, given the environment of defectors in which they found themselves, cooperators' decisions to withdraw cooperation over time were wholly rational. Figure 8, above (p.157), shows a strikingly similar pattern to that seen in a classic adult public goods game (see, for example, Fehr & Gächter, 2002). Thus, this finding adds to a new but

growing body of evidence that humans are skilled conditional cooperators from at least as young as their late pre-school years (Lergetporer et al., 2014; House et al., 2013b; Vogelsang et al., 2014). More remarkable, as touched upon above, is that they also appear to exhibit conditional cooperation in a chicken game. In this context, doing so is wholly pig-headed and counterproductive, and thus suggestive of a large degree of spitefully-motivated hyper-competitiveness.

7.7.3: Age Effects

In the stag hunt experiments, no monotonic age effects were found. The same interpretation of the stag hunt appeared to be applied by a majority of children throughout middle childhood. This in itself is not unprecedented. Although there is a broad tendency for prosociality to increase with age (Eisenberg, Fabes & Spinrad, 1998, Fehr et al, 2008), some evidence suggests that this process appears to stall during middle childhood (House et al, 2013), and, perhaps consequently, prominent previous game theoretical studies have sometimes failed to find significant age differences during this window (e.g. Lergetporer et al., 2014, Sally & Hill, 2006). It seems likely that if the sample in this chapter's stag hunt experiments had been extended to include larger numbers of children older than 8-years, there would have at some point been a marked increase in Stags, but at exactly what point this switch occurs remains a topic for future research.

More surprising is the age effect observed in Experiment 5d, the chicken game. Not only were the youngest children consistently the most cooperative, but amongst the eldest children, those aged 9+, cooperation declined and did so in a context whereby this action was materially counterproductive in terms of payoffs. The finding that older children's decisions deviated more from the SEM than those of younger children is not unusual. What is strange is seeing this occur in a mutually destructive context, an outcome which is hard to reconcile with the presence of either norms or fairness preferences, both of which should be more developed in those in mid-to-late childhood than in younger children. Instead, in the chicken game, the desire to win overcame both.

This appears to be strong evidence in favour of the presence of hyper-competitiveness. Given the growing frequency of defections as the game wore on, many participants likely inferred that defection would remain the most likely move on their opponent's part going forward, yet they still largely defected. It seems many preferred to spitefully bring about the "Crash"

scenario than allow a relative advantage to their opponent. The willingness to pay a cost to avoid relative disadvantage in children up to around 10-years of age has been observed before, but not in such a stark and repetitive form as was seen here.

7.7.4: Relative Evaluation and Hyper-Competitiveness

There is growing evidence that children in early to middle childhood are willing to violate the SEM by paying to reduce a partner's payoffs. They will often reject outcomes which are weighted in a peer's favour and will willingly inflict absolute costs on themselves if it means not being relatively worse off (Bereby-Meyer & Fiks, 2013; Blake & McAuliffe, 2011; Bügelmayer & Spiess, 2014; McAuliffe et al., 2013; McAuliffe et al., 2014; Sheskin, Bloom, et al., 2014). Importantly, there is a definite social element to this. Specifically, children are known to be more likely to pay to reject numerical disparity when a peer is involved than when the entity with more than they have is a random number generator or simply an empty space (Bereby-Meyer & Fiks, 2013; McAuliffe et al., 2013).

What the present chapter appears to show, however, is something beyond this: children paying an absolute cost not to *avoid* disadvantageous inequity, but instead to actively *acquire* advantageous inequity. In other words, the hyper-competitive motivation at play here was not to avoid an upwards social comparison to a richer neighbour by moving from relative poverty to a position of equity, but instead to generate a downward social comparison towards a partner who was previously one's equal, even if this meant impoverishing both parties in the process. Anecdotally, many of the participants, when quizzed informally after the experiment, made comments which further support this interpretation. Several of those who spent the latter rounds of the experiment stuck in the 1-1 Hare equilibrium, when asked as to whether they would have preferred a 3-3 outcome or one weighted 1-0 in their favour, enthusiastically stated that the latter would have been preferable. I know of only one other recent study to have investigated this possibility and reached the same conclusion. Sheskin et al (2014) found a similar effect, albeit only in those aged 5-6 years, and suggested that young children's ORPs are dominated by an unusually strong sense of social comparison, which they need to overcome before they can begin to prefer fair outcomes. The results of Experiments 5a-5d partially support this conclusion, although here the hyper-competitive behaviour extended to children older than 5-6, something Sheskin et al did not find. Clearly, more work needs to be done towards investigating the extent and limitations of this phenomenon, and this a theme which will be returned to in the next chapter. For now, it is sufficient to say that the work here

provides the single strongest piece of evidence for hyper-competiveness in pre-schoolers unearthed to date.

7.7.5: Comparisons with previous stag hunts

The results presented in this Chapter are very different to those seen in previous stag hunts with children (Duguid et al., 2014; Wyman et al., 2013). How can this disparity be explained? Whilst generally I feel that my results provide good evidence in favour of strong hyper-competitiveness in early-to-middle childhood, it should be acknowledged that aspects of the methodology employed may have partly inhibited cooperativeness.

In previous SH experiments, the set-up was less explicitly game theoretical. Instead of making binary choices which generated numerical responses, participants – aged 4-years old, the same as the youngest participants in the present set of experiments – made physical movements to operate an apparatus. Stag and hare outcomes were differentiated by, respectively, high and low value stickers (Wyman et al., 2013), and desirable and mediocre food rewards (Duguid et al., 2014). By contrast, although my instructions carefully omitted any references to competitive terms such as “game”, “win” or “lose”, it is true that any paradigm featuring “points”, and “prizes” accrued by playing with sets of cards perhaps has an intuitively game-like schema, and may have been interpreted thus by participants. Whilst, as noted, the use of points or their equivalent as a form of pseudo-currency is widespread in developmental game theory tasks, it may be the case that playing for visible, physical prizes such as food would have resulted in less inhibitory restraint in rejecting larger absolute rewards. Investigating whether and to what extent this is the case would be a useful avenue of future research, and of interest to all who investigate developmental economics.

Even more so than this, however, the most crucial source of the disparity between the stag hunts presented here and previous efforts is probably the degree of communication allowed. In the experiments in this chapter, the children could clearly see each other, but were not encouraged to verbally communicate. In the experiments of Wyman et al, children were either given a verbal and physical hint by a confederate (Wyman et al., 2013), or were free to talk amongst themselves (Duguid et al., 2014). The fact they were able to do this seems to have been a major factor in allowing them to outperform chimpanzees in the latter study. From experiments 5a-5d, however, it would seem that verbal communication may be vital not

just in allowing children to outperform chimpanzees, but in allowing them to succeed in the experiment full-stop.

7.7.6: Conclusion

Certainly, in the absence of communication subjects did not seem to intrinsically value a mutualistic outcome for its own sake. Instead, victory appeared to be the goal of many of the children in the experiments reported in this chapter. This may have been misguided in rational, material terms, but children in these experiments seemed to have placed more utility on achieving relatively more than their partner than they did to achieving a mutualistic outcome. The presence of these hyper-competitive children consequently caused the more cooperative children in the sample to withdraw their cooperation conditionally. More work will clearly need to be done on the SH and on mutualism in children generally, but this chapter shows that it is certainly not a given assumption that children prefer mutualistic cooperation under all conditions, or that they automatically value it intrinsically. Exactly what is necessary and sufficient for mutualism to emerge will be an interesting topic of future research. In the absence of communication, however, it seems that many children aged 4-8 years actually prefer relative advantage over a mutualistic outcome, and do not mind if they have to pay to achieve this.

Chapter 8: General Discussion

The aim of this dissertation was to examine the emergence and development of hyper-competitive behaviours in childhood, by focusing on negative other-regarding preferences (ORPs), and how children's decisions consequently differed from those predicted by the standard economic model (SEM). In experiments 1-3, deviation from the SEM was signified by second party punishment. In experiment 3 this was compared and contrasted with rewards. In experiment 4-5d, deviation from the SEM consisted of a refusal to cooperate. Collectively, the experiments presented in this dissertation set out to shed light upon the following four areas of particular interest to the study of negative ORPs[10]:

(1) How does the strength of negative ORPs change across the breadth of childhood? Does it increase or decrease? Is this process linear, and does it occur reliably across different experiments?
(2) How do negative ORPs compare in this respect to positive ORPs? Do they develop in the same way, at the same time and at the same pace? Is a tendency to violate the SEM something we simply learn with age, and are all violations of the SEM driven by the same preoccupation with fairness?
(3) What can the expression of ORPs in childhood tell us about the developing understanding of fairness preferences and norms?
(4) What are the implications of negative ORPs for children's ability to cooperate? Negative ORPs tend to be looked at in isolation, i.e. in experiments where punishment is the focus and to punish or not to punish is the outcome variable. How do negative ORPs affect children's behaviours in other contexts?

The findings presented here suggest that negative ORPs are a strong motivator of children's behaviour throughout their pre-adolescent years. In this respect they appear to take root at least as strongly as positive ORPs, a tendency which can hinder many children's abilities to achieve cooperative outcomes. This discussion will first review the collective findings from all eight experiments, and how they compare to, and extend, previous findings in the literature. It

[10] Analysis of all four of these issues will be provided throughout the following discussion. In particular, however, in depth focus relevant to each specific area is particularly concentrated in the following sections. For area (1), see section 8.2.1; for area (2), see sections 8.1.1., 8.2.1. and 8.2.3; for area (3), see sections 8.2.2. and 8.3.3; and for area (4) the reader is directed to sections 8.1.2, 8.2.2, and 8.2.5.

will then go on to explore some of the theoretical and methodological implications for the topics of children's social behaviours and other-regarding preferences.

8.1: Comparing the Findings of this Dissertation with Previous Studies

Existing evidence on negative other-regarding preferences and hyper-competitiveness in children is relatively sparse, so this dissertation is in many ways exploratory. Some of the findings presented here chime well with what was known already. Others appear much more challenging to the orthodox view of children's social behaviour.

8.1.1: The Reciprocity Studies

Mini-Ultimatum game with theft: (negative) reciprocity

In Chapters 3 and 4, both using the mini-ultimatum game with theft, the focus was on costly punishment. The prevalence of costly punishment increased with age, and a particularly pronounced shift occurred between children of pre-school age, whose behaviour often did not greatly differ from the SEM, and those from around 6-years upwards. Chapter 4 also demonstrated that most children throughout pre-adolescence did not apparently attend to the distinction between the outcome of a proposer's offer and the proposer's intention. This suggests that spiteful, hyper-competitive behaviour during this phase is probably motivated chiefly by inequity aversion rather than moral or normative considerations of another's motives. Another novel finding from this pair of experiments was the behaviour of proposers, which was unusually stingy compared to that shown in previous experiments.

The observed increase in spiteful punishment with age corresponds with previous work which shows this pattern both in developmental ultimatum games specifically (Harbaugh et al., 2003; Sally & Hill, 2006; Steinbeis et al., 2012), and in other conceptually similar experiments in which children must pay to reject inequity (Blake & McAuliffe, 2011; McAuliffe et al., 2013; McAuliffe et al., 2014). These experiments extend such findings by demonstrating this effect across a wider age spectrum. They also replicate and extend what a number of recent studies have either hinted at (Takagishi et al., 2014) or tested directly (Bereby-Meyer & Fiks, 2013; Wittig et al., 2013), namely that in the ultimatum game responders' behaviour barely, if at all, differs from the SEM until responders reach at least 6-years of age. Indeed, punishment of all

types, both costly and non-costly, increased over the course of childhood, suggesting that rising punitive sentiment generally is behind this, rather than an ability to overcome selfish tendencies and violate the SEM.

Experiment 2 also sought to add clarity to the debate as to whether child responders differentiate between outcome and intention when enacting costly punishment. In common with some (Gummerum & Chu, 2014; Güroğlu et al., 2009), though by no means all (Bereby-Meyer & Fiks, 2013; Sutter, 2007) of the previous work on this topic, I found that this was something which most pre-adolescents do not do. This they have in common with chimpanzees (Jensen et al., 2007a; Kaiser et al., 2012).

One novel finding of Experiments 1 and 2 is that proposers of all ages had little if any preference for fair outcomes, and no obvious aversion to advantageous inequity. This suggests that framing effects, in this case in the form of demand characteristics favouring theft rather than giving, can disrupt previously observed preferences for equitable outcomes from mid-childhood upwards (Blake & McAuliffe, 2011; McAuliffe et al., 2013; McAuliffe et al., 2014).

The Moonlighting game: (negative and positive) reciprocity

Chapter 5 extended this focus on costly second party punishment (or negative reciprocity) by placing it alongside costly positive reciprocity in a directly comparable manner, using the first adaptation of the moonlighting game for a developmental sample. The results unequivocally demonstrated that negative reciprocity was the stronger motivation throughout pre-adolescence and that, given a directly comparable choice, child subjects are much more likely to violate the SEM in order to act upon negative other-regarding concerns than positive other-regarding concerns. There was also no evidence that positive and negative reciprocity are linked personality traits within-subjects, suggesting that they are not simply mirror-image expressions of a single cognitive preoccupation with fairness. Another intriguing finding from this experiment regarding negative reciprocity was that it did not appear to be particularly effective amongst children, since theft rates actually increased as the experiment progressed.

Previous attempts to compare positive and negative reciprocity in childhood have tended not to do so under conditions of equivalent cost (Keil, 1986), or have focused only on children ages 11-years and above (van den Bos et al., 2012). Nevertheless, after correcting for these aspects, my results are still largely complementary to those of both of these studies. Children are negatively reciprocal more than they are positive reciprocal. They are more willing to spitefully

punish than they are to enact reciprocal altruism. Interestingly, this may also be true of adults, who are fairly consistent in their punishment of miserly ultimatum game offers (Camerer, 2003; Gächter, 2014), but of only limited trustworthiness in the trust game (Berg et al., 1995; Sutter & Kocher, 2007). Importantly, in the original moonlighting game experiment, adults too were more consistent negative than positive reciprocators under conditions of equivalent cost (Abbink et al., 2000), a finding that has since been replicated in a similar experiment (Keysar et al., 2008). This finding adds weight to similar conclusions reached by experiments and questionnaire studies with adults (Dohmen et al., 2008; Keysar et al., 2008; Perugini et al., 2003; Peysakhovich et al., 2014; Yamagishi et al., 2012) by showing that not only are the two not expressed to co-varying degrees by adult participants, but they also do not appear to emerge in a similar way in ontogeny. This in turn suggests that they may not be the product of a single sense of reciprocal justice.

This does not mean that the overall implications of strong reciprocity theory are incorrect, since at an aggregate level both reciprocity behaviours are common with the population. However, it does mean that a simple preoccupation with "reciprocity" is not a common human trait, an assumption made by strong reciprocity theory, and also prominent models of fairness concerns and norm psychology (Bolton & Ockenfels, 2000; Gintis, 2000b; Haidt & Joseph, 2004). Instead, the proclivities for positive and negative reciprocity are distinct personality traits. Of the two, negative reciprocity seems to be in place more strongly at an earlier stage of ontogeny. Our desire to punish free-riding seems to occur before our willingness to engage in widespread rewarding behaviour or indeed, if the work in Chapters 6 and 7 is to be believed, in risky acts of cooperation. In terms of reciprocity, we develop a taste for hyper-competitiveness before hyper-cooperativeness.

8.1.2: The Cooperation Studies

The Battle of the Sexes Game

Experiment 4 applied the battle of the sexes game to a developmental sample for the first time, in order to examine if and how hyper-competitive motives would impact on actual cooperation. It was found that an unwillingness to incur disadvantageous inequity in the short-term greatly restricted the efficiency of subjects' overall performances. Despite explicit instructions to treat the task as a positive-sum exercise, most subjects seemed unable to

follow this advice, and instead preferred outcomes which reduced their own scores but also eliminated negative inequity between their own score and that of their partner's.

The levels of cooperation observed in Experiments 4-5d were surprisingly low. Of the two studies, the findings from the battle of the sexes game (BOS; see Chapter 6) are probably easier to reconcile with previous findings. Throughout pre-adolescence, children have previously been observed preferring advantageous inequity (Fehr et al., 2008; Fehr et al., 2013; Sheskin, Bloom, et al., 2014), even if they become increasingly skilled at hiding this fact, thus displaying a growing awareness that this is not socially acceptable behaviour (A. Shaw et al., 2014). It is true, however, that not all studies have observed this preference for advantageous inequity, participially amongst children aged 8-years and above (Blake & McAuliffe, 2011; McAuliffe et al., 2013; McAuliffe et al., 2014). Very consistent across all previous findings, however, is dislike of disadvantageous inequity and a willingness to pay to avoid it. In Chapter 6 I extend these findings by showing that this preference is not limited to one-off encounters, but is strong enough to cause persistently intransigent behaviour which prevents cooperation from occurring in a potentially mutualistic context, to the cost of everyone involved. Importantly, the results demonstrate that this appears to be down to inequity aversion and relative evaluation, and not due to risk aversion or uncertainty, since altering the level of both of these factors did not affect the willingness of children to cooperate at any age bracket. There did, however, appear to be some degree of reinforcement learning towards avoiding risky scenarios (see Chapter 6 Discussion, Section 6.4).

The Stag Hunt

Experiments 5a-5d focused on a younger sample, and looked at how hyper-competitive motivations might affect cooperation in a more unambiguously mutualistic context, the stag hunt. Children largely failed to achieve stable, mutual cooperation in the stag hunt, even when paired with a helpful confederate. When paired with peers, children's failure to consistently cooperate caused a pattern of conditional cooperation. The reason behind this appeared to be that many of them preferred to try and "win" the stag hunt game by gaining advantageous inequity in preference to maximising their own absolute payoff. A follow up experiment using the chicken game added weight to this latter argument by showing that children tended towards greedy, risky play rather than modest, conservative play in a similar game theoretical paradigm, thus suggesting that their performances in the stag hunt were not due to risk aversion.

The results of these four experiments are not consistent with previous findings in the wider literature, although they are consistent with those of the other studies in this dissertation. Pre-schoolers, and even toddlers, have previously demonstrated many advanced cooperative abilities (Svetlova et al., 2010; Tomasello, 2014) and, crucially, 4-year old children have previously been observed to succeed in cooperating at conceptually very similar tasks (Duguid et al., 2014; Grueneisen et al., 2015; Wyman et al., 2013). Nevertheless, in Experiments 5a-5d children of this age, and indeed elder children aged up to 8-years, failed to achieve consistent cooperative outcomes. The methodological differences between my study and those of Wyman and colleagues almost certainly played a part in this outcome, and these are discussed in Chapter 7 (Section 7.7.5). There is little precedent for the present stag hunt findings in the experimental literature, but clearly more work needs to be done on children's attitudes and preferences towards mutualistic cooperation. Some avenues for future research on this topic are suggested in Section 8.4 (below).

One interesting finding from Chapter 7 which is consistent with the majority of existing work is the demonstration of conditional cooperation amongst children, something which until recently was thought not to occur (Harbaugh & Krause, 2000), but is now reliably beginning to appear across a range of different economic experiments (Alencar et al., 2008; Angerer et al., 2015; Lergetporer et al., 2014; Vogelsang et al., 2014). Children, it would seem, are conditional cooperators. Work looking at their ability to target strategic cooperation appropriately had previously shown that this begins at around 5-6 years (House et al., 2013; Steelandt, Dufour, Broihanne, & Thierry, 2012), but Chapter 7, in particular in study 5c, suggests that this ability may be present from an even earlier age.

8.2: Theoretical Considerations

8.2.1: Age effects

An important consideration when comparing my findings to those of the wider literature is the relationship between increasing age and violations of the SEM, both positive and negative. Very broadly, adults are more pro-social than children, and consequently if and when linear trends in prosocial behaviour are observed they tend to show children becoming more hyper-cooperative the closer they get to adulthood (Eisenberg et al., 1998). The present dissertation

has shown that the same may not be true of negative ORPs and spiteful violations of the SEM, or at least not in a way which generalises to all behavioural contexts.

Violations of the SEM through Second Party Punishment

Certain hyper-competitive behaviours, such as costly punishment, did increase with age more often than not. Costly punishment was observed reliably increasing with age on multiple occasions in Chapters 3 and 4 (MUGWT), though the same did not occur in a sample with an older mean age in Chapter 5 (moonlighting). However, it should be pointed out that Experiments 1 and 2 suggest that the greatest shift from adherence to the SEM towards costly second party punishment occurs at around 6-7 years of age. Experiment 3 did not test children younger than this, thus perhaps explaining the lack of linear age effects.

One thing which did not occur was any unexpected linear reduction in punitive sentiment in conjunction with increasing age. Indeed, to my knowledge such a finding has never been observed in a developmental experiment. Collectively Experiments 1-3 suggest that if and when age effects on the level of costly punishment occur during pre-adolescence, they will tend to be positively related. Data from previous experiments with teenagers, however, suggest that this relationship may change once adolescence is reached, upon which more complex considerations such as the outcome/intention distinction come into play (Gummerum & Chu, 2014; Güroğlu et al., 2009; Sutter, 2007).

Violations of the SEM through Non-Cooperation

Costly punishment is not the only means through which negative ORPs can be observed. Experiments 4-5d showed this clearly, by demonstrating that children frequently refused to cooperate when doing so would present them with an absolute gain combined with a relative disadvantage, something I will henceforth refer to as “zero-sum thinking”. Additionally, Experiments 5a-5d appeared to suggest that many young children went even further, preferring to have more than their partner than to cooperate mutualistically, something which can be deemed “seeking relative advantage”. Both behaviours are clearly hyper-competitive, and violate the SEM because they require the actor to pay a (rationally needless) cost in order to reduce another’s income.

A key finding of this dissertation is that both of these motivations simply seemed to be present to a high and fairly inflexible degree across the entire spectrum of pre-adolescence. There

were no ontogenetic changes. This appears to be a novel finding, although other studies hinting at both of these phenomena in early-to-middle childhood have recently begun to emerge (Häger, 2010; Sheskin, Bloom, et al., 2014). Importantly, the presence of these motivations means that hyper-competitive behaviour may be expressed in a less linear way across childhood than hyper-cooperative behaviour. The latter, on the whole, has to be learned, and thus tends to increase as children get older (Benenson et al., 2007; Evans et al., 2013; Fehr et al., 2008; Fehr et al., 2013; Gummerum et al., 2010; Häger, 2010; House et al., 2013; Sutter & Kocher, 2007). In addition to genuine altruism, increasing positive violations of the SEM with age can probably also be attributed to two other factors. These are greater internalisation of social norms (Rutland et al., 2005; Smith et al., 2013), and greater awareness of the importance of maintaining a good reputation (Engelmann et al., 2012; Engelmann et al., 2013).

Virtually all expressions of hyper-cooperativeness require the selfish motivation to maximise one's income to be overcome. This change from egotism to altruism is generally normative – it is hard if not impossible to think of examples where preferring the latter to the former would not be considered "good" behaviour. Younger children often struggle to meet such standards compared to older children, however, due to their greater selfishness (Benenson et al., 2007; Gummerum & Chu, 2014; Gummerum et al., 2010; Gummerum et al., 2009; Smith et al., 2013).

Hyper-competitive behaviour may be more complex than this. Some acts of behaviourally spiteful, negative other-regarding behaviour, are similarly normative – third party punishment is a good example (Gummerum & Chu, 2014; McAuliffe et al., 2015; Riedl et al., 2015; Robbins & Rochat, 2011), although second party punishment can also be motivated at least in part by normative concerns (Kurzban et al., 2015). The tendencies to act against wrongdoers and arbitrate fairly in disputes are exactly the types of behaviours which children are socialised into, and therefore probably increase in frequency over the course of ontogeny. The punishment in this case, and thus the outcome of the negative ORP involved if not the actual thought process behind it, can be technically "altruistic", in the sense that the punisher often does not reap a net material benefit from enforcing normative or cooperative behaviour on behalf of a third party, or of the social group as a whole (Fehr & Gächter, 2002).

Other expressions of hyper-competitive behaviour are less normative, however, in a way for which it is tricky to imagine a hyper-cooperative equivalent. The zero-sum thinking and relative advantage seeking in Experiments 4-5d provide perfect examples. These are not the types of behaviours children are generally encouraged to engage in – quite the opposite, in fact.

Consequently, these types of behaviours probably decrease in frequency during ontogeny, although the results presented in Experiments 4 (BOS) and 5d (Chicken Game) in particular suggest that this decrease may not occur to a high degree until adolescence. Other recent work also hints that this may be the case, possibly because emotions such as envy and schadenfreude are present in higher levels in children than is the case in adolescents (Fehr et al., 2013; Sheskin, Bloom, et al., 2014; Steinbeis & Singer, 2013). A decline in such sentiments would explain a reduction in the strength of disadvantageous inequity aversion and relative advantage seeking respectively.

If this is true then it means that the conventional division of ORPs into "altruistic" and "spiteful" (e.g. Levine, 1998) is insufficient. Hyper-cooperative, functionally spiteful behaviour can be expressed in two distinct ways. One, concerning justice and the enforcement of fairness, is normative and therefore in short supply amongst young children. It largely needs to be socially learned, and thus increases across ontogeny. The other, concerning zero-sum thinking and relative advantage seeking is not normative, indeed quite the opposite, and yet appears to be very commonly expressed in the behaviour of young children. Consequently, for this type of hyper-competitiveness, the behaviour which needs to be learned is to *not* act upon one's negative ORPs. This means that negative violations of the SEM may not have a monotonic relationship with age. Instead, there are two subtypes, normative and non-normative hyper-competitiveness, which develop in opposite directions. Spiteful violations of the SEM are thus not necessarily as likely to have a positive relationship with ontogeny as positive violations. One thing which all violations of the SEM do appear to share, however, is that they are motivated to a large degree by fairness concerns.

8.2.2: Evidence of Fairness Concerns

There are various proposed universal human fairness concerns, and on the whole they are complementary to each other rather than conflicting. This dissertation has provided clear evidence that they do not occur in behaviour at the same point in ontogeny. Indeed, some were abundant throughout childhood, whilst others seemed almost totally absent, raising the question of when in development they do occur.

Inequity Aversion

People are inequity averse, meaning that they will pay a cost to avoid an unequal distribution which is stacked against them (disadvantageous inequity aversion; DIA) and sometimes also when it is in their favour (advantageous inequity aversion; AIA) (Fehr & Schmidt, 1999; Loewenstein et al., 1989). There was plenty of evidence of DIA throughout all experiments, an expected finding, since children as young as 3-years show evidence of this (LoBue et al., 2011), and it is generally thought to steadily increase from this age until at least 8-years (Fehr et al., 2008; McAuliffe et al., 2013; McAuliffe et al., 2014; McAuliffe et al., 2015). Contrastingly, however, there was little AIA in evidence in any of the experiments, as witnessed by the extremely high levels of theft in the MUGWT and moonlighting experiments, and the consistent willingness of participants to defect in the two mutualistic cooperation experiments.

It has been previously suggested that advantageous inequity aversion becomes a strong preference in children at around 8-years of age (Fehr et al., 2008; McAuliffe et al., 2013; McAuliffe et al., 2014; McAuliffe et al., 2015). It should be noted that this finding is not always replicated, and other-experiments have shown it to be in short supply well into late childhood, and even the teenage years (Fehr et al., 2013; A. Shaw et al., 2014). It is true, however, that existing theories of inequity aversion are quite clear that, even amongst adults, AIA is a weaker motivation than is DIA (Fehr & Schmidt, 1999; Loewenstein et al., 1989). A tolerance amongst the majority for relative disadvantage does not seem to emerge until adolescence (Fehr et al., 2013; Sheskin, Bloom, et al., 2014) a suggestion which is supported by the consistent finding throughout this dissertation that very few subjects seemed willing to tolerate it in a wide variety of contexts across all eight of the experiments.

Advantageous inequity aversion also appears to be easily disrupted amongst adults. Experiments which, like the mini-ultimatum game with theft, frame a distribution task as theft as opposed to giving can elicit unusually greedy behaviour amongst adult subjects too, causing them to repeatedly choose advantageous inequity instead of a fair outcome, thus showing little or no AIA (Bosman, Sutter, & van Winden, 2005; Bosman & Van Winden, 2000, 2002; Keysar et al., 2008). With this in mind, perhaps the children in the present studies should not be judged to harshly for succumbing to similar temptations. What such findings do show, however, is the fragility of AIA to context, and the fact that it does not simply become an entrenched preference amongst all normal people once middle childhood is reached.

Equity, Reciprocity, Competition (the ERC model)

To Fehr and Schmidt's (1999) concerns with equity, Bolton and Ockenfels (2000) add a preoccupation with relative standing and a concern for reciprocity. The evidence for these preferences in childhood is similarly mixed. Evidence of a concern for relative standing is extremely widespread in this dissertation, being in strong evidence in the behaviour of children in all eight experiments. Indeed, it would be no exaggeration to say that the evidence presented here, combined with other recent work (Blake et al., 2014; Bügelmayer & Spiess, 2014; Fehr et al., 2008; House et al., 2013; Sheskin, Bloom, et al., 2014; Steinbeis & Singer, 2013) shows that pre-adolescent children appear to be very strongly fixated upon it. Attempting the "win" a stag hunt is particularly good evidence for this, as is the general tendency in favour of "seeking relative advantage" discussed in the previous section. This is particularly striking when one considers that other primates are indifferent to whether others have an equal or lesser amount to them, so long as they are themselves rewarded (Burkart et al., 2014; Jensen et al., 2006).

The strength of this emphasis on our relative as opposed to our absolute outcome appears to be a curiously human phenomenon, and is excellent evidence of one of our strongest and earliest hyper-competitive preferences. An important point to add here is that a minority of subjects retain preferences such as zero-sum thinking and relative advantage seeking into adulthood. For example, around 10-20% of adult subjects (depending on the degree of anonymity allowed) are willing to spitefully pay to convert an offer from the experimenter of an even endowment for themselves and a partner to one which is skewed in their favour (Abbink & Herrmann, 2011). Furthermore, in the "money-burning game" in which, as the name suggests, subjects can destroy each other's endowments, those at the top are disproportionately targeted for impoverishment by those with less (Zizzo & Oswald, 2001). This suggests that although a preoccupation with relative standing may be mitigated in older children and adults by greater socialisation and more benign fairness concerns (e.g. social welfare), it does not disappear. Echoes of this type of decidedly non-normative, spiteful violation of the SEM can also be seen in the literature on anti-social punishment (Herrmann et al., 2008; Sylwester et al., 2013) discussed in the introduction (see Chapter 1, Section 10).

On reciprocity the picture is more mixed. Certainly, costly negative reciprocity is common behaviour in children by the time they are six years old, though it does not necessarily appear to be common behaviour prior to this. Positive reciprocity was far from negligible in Experiment 3 (occurring on around 35% of possible trials), but it could not be considered preferred behaviour, with most responders choosing defection over reciprocation. An interesting addendum to this, however, is that negative and positive reciprocity do not appear

to constitute two sides of a single fairness concern, as shown by Experiment 3. Children's preference for positive reciprocity is weaker than their preference for its punitive equivalent, and this may also be true of adults (Abbink et al., 2000; Keysar et al., 2008). There certainly seems to be little direct correlation between the two preferences within-subjects (Dohmen et al., 2008; Perugini et al., 2003).

Maximising Social Welfare

The third classic source of theories of fairness concerns is the work of Charness and Rabin (2002), who further added conditional cooperation[11] and a preference amongst some subjects for "increasing social welfare" by paying a small cost to maximise net outcomes for the group. The presence of conditional cooperation in Chapter 7 and in other studies has already been discussed. Children appear to be adept conditional cooperators by as early as 4-5 years of age. The second of Charness and Rabin's fairness concerns, increasing social welfare, was conspicuous by its absence throughout the present dissertation. In particular, in the BOS and stag hunt, pre-adolescent children consistently failed to choose the option which would have maximised social welfare even when doing so conferred upon them an actual *gain*, so it seems very unlikely that they would have paid a cost to do so, although this remains an untested empirical question for now. It is perhaps possible that some children would be prepared to pay to increase another's income if they were nevertheless able to retain a relative advantage for themselves whilst doing so.

One existing study which can potentially shed some light on children's tendencies to maximise social welfare is that of Shaw and Olson (2014). Here, the researchers found that when dividing resources (in this case five erasers) amongst two strangers, 6-8 year olds would rather throw away the spare eraser than give it to one of the strangers. In other words, subjects considered two erasers each and the disposing of the fifth to be a superior outcome to giving three erasers to one child and two to the other. This is good concurrent evidence that social welfare maximisation, even when it costs the actor nothing, is not something that seems to readily occur to many children. Instead, in this example, a preference for equal outcomes took precedent even in a scenario in which the inequitable alternative would only have affected third parties and not the actual subject of the experiment. In an equivalent situation in which adult subjects were asked to split five $10 bills between two other people, it seems extremely unlikely that a majority would decide that handing out two each and then destroying the fifth

[11] Or, as they put it, a tendency to "withdraw willingness to sacrifice to achieve a fair outcome when others are themselves unwilling to sacrifice" (p.817).

was an optimal outcome. Maximising social welfare, it seems, may be one of the most complex fairness concerns, and one which does not emerge until late childhood or adolescence. As the results of experiments with adult subjects such as the money burning game show, however, it would seem that some people never acquire the desire to maximise social welfare, or at the very least that they often fail to apply it (Abbink & Herrmann, 2011; Zizzo & Oswald, 2001).

Evaluating Outcome and Intention

Like social welfare, the ability to evaluate and distinguish an external actor's intentions towards oneself from the actual outcome they caused (Falk et al., 2003, 2008; Nelson, 2002), or at least to act upon the former and not the latter, seems to be one of the more complex and sophisticated fairness concerns. The distinction between outcome and intention is a subtle one, and one which reaction time data shows that even adults require markedly more time to weigh up when acting as responders (Gummerum & Chu, 2014), suggesting that they too find it a relatively challenging judgement to make. It may be that most people are unable to reliably make it until they have experienced the final maturation of their intentionality and mind-reading abilities in late adolescence (Blakemore et al., 2007; Dumontheil et al., 2010). Additionally, however, based on the findings of spiteful disadvantageous inequity aversion in Chapters 3-4, as well as in many other previous studies (Blake & McAuliffe, 2011; Fehr et al., 2008; Fehr et al., 2013; McAuliffe et al., 2013; McAuliffe et al., 2014; Sheskin, Bloom, et al., 2014), there is an alternative possibility. It may also be the case that many children *understand* the outcome/intention distinction reasonably well, but are nevertheless unwilling to agree to a comparatively generous offer which still presents them with a relative disadvantage. In other words, they may recognise that sometimes 25% can be a generous offer, and yet nevertheless cannot overcome their aversion to inequity enough to gratefully accept it.

Children's attendance to Fairness Concerns

Cumulatively, the evidence shows that fairness concerns are an important source of children's negative ORPs. They can provide considerable explanatory power to the results of Experiments 1-5d. Much of the competitive and hyper-competitive behaviour on display can be explained in terms of fairness concerns such as disadvantageous inequity aversion, relative evaluation, and to some extent costly negative reciprocity. Perhaps one of the reasons hyper-competitive behaviours are so pronounced amongst children is that more sophisticated, some might even say moralistic, fairness concerns such as social welfare, advantageous inequity aversion, outcome/intention differentiation and costly positive reciprocity are not present to as high an

extent in pre-adolescents as are the former, more basic set. This may be because these more altruistically-motivated fairness concerns, which are generally driven by positive rather than negative ORPs, require considerable social learning and internalisation of social norms before they become entrenched behavioural preferences.

8.2.3: Evidence of Social Norms

There was little evidence of social norms acting as a source of children's ORPs in this dissertation. This is in spite of considerable evidence from the literature suggesting that they might have been expected to do so. Children develop an understanding of social norms from as young as 3-years old (Rakoczy & Schmidt, 2013; Rakoczy et al., 2008; Rossano et al., 2011; Smith et al., 2013), an ability so prodigious that it has been plausibly suggested that *Homo sapiens*' minds have evolved to naturally attend to and understand them (Chudek & Henrich, 2011; Rakoczy & Schmidt, 2013; Tomasello, 2009).

It is true that in the sphere of negative ORPs specifically, it is not easy to precisely determine when acts such as second party punishment are motivated by norms such as reciprocity or fairness. Other motivations like strategic attempts to incentivise future cooperation, inequity aversion, or even just simple vengeance, are hard to rule out (Kurzban et al., 2015; Raihani & McAuliffe, 2012). If there is any evidence at all of subjects' norms affecting my findings on negative ORPs, it can perhaps be found in the lack of a strong preference for punishing offers of zero amongst responders in the pre-school age cohort. This act is clearly unfair, and yet the sample of pre-schoolers as a whole did not actively prefer to punish it, even when doing so would not have required a violation of the SEM. Indeed, once it became apparent to responders aged around 4-5 years that they could not accrue any points for themselves, many of them seemed surprisingly indifferent to the prospect of punishment. Even here, however, the more punitive responses from middle childhood upwards may have been motivated by a stronger sense of DIA to at least as high a degree as normativity.

Better evidence for a lack of subjects' consideration of social norms is the relative absence of positive ORPs across multiple studies. Proposers repeatedly violated the fairness norm by turning an equal offer into an unequal offer in the MUGWT. Responders repeatedly violated the (positive) reciprocity norm by not returning gifts in the moonlighting game. Very large percentages of children were reluctant to fulfil norms relating to cooperating, sharing and helping in the BOS and the stag hunt.

So, what went wrong? Over the course of my eight experiments, there was little evidence of normative behaviour, aside from some modest improvements from those in middle- to late-childhood in Experiments 1 and 2. The Shaw and Olson (2014) eraser study described in the previous section shows clear evidence of children aged 6-8 years thinking about and applying fairness norms, albeit in slightly idiosyncratic way. Why, then, were subjects seemingly so reluctant to do so in the experiments reported in the present dissertation? One key reason was that in the Shaw and Olson study children were acting as detached arbiters of procedural justice. The resources on offer were not their own, and nor was there any way that they could acquire them. There is little doubt that when the potential to selfishly acquire resources is removed from the experimental procedure like this, children's ability to apply fairness norms comes to the fore to a much greater degree. Under such circumstances, pre-school children have been observed verbally admonishing norm violators (Rakoczy et al., 2008; Rossano et al., 2011), and even enacting third party punishment (Kenward & Östh, 2012; Riedl et al., 2015; Robbins & Rochat, 2011), a highly sophisticated social behaviour.

Children, then, do care about norms from early in life. The problem is that they also care about self-maximising, and the temptation to do the latter often overrides their ability to do what they know is the right thing. In the terminology of the norm psychology literature, young children are therefore capable of obeying norms "instrumentally" – meaning that they understand them, and will obey them if the stakes are low or if compelled to – but have not truly internalised them to the point that they derive actual satisfaction from observing them themselves (Gintis, 2007).

Haidt's (2001) thoughts on the ontogeny of normative internalisation are of interest here. He argues that the average developing child will, all other things being equal, naturally reach a stage of development where they are willing to truly internalise social norms somewhere between late childhood and early-to-mid adolescence. This idea has not been extensively explored, but several recent studies are consistent with it. A recent dictator and ultimatum game study by Kogut (2012) found that, although children aged 7-8 years would make largely fair distributions, they reported feeling unhappy about this afterwards, suggesting that they were reluctant to do so. By 9-10 years, however, children both shared equally and reported feeling good about this. This suggests that by this stage greater normative internalisation was in evidence, in that sharing had become a cognitively rewarding act, and thus an end in and of itself by causing them to feel the famous "warm glow" of moral self-satisfaction (Andreoni, 1995).

Similarly, a study on children's attitude to the norm of not expressing ethnic prejudice found that virtually all-children in a sample aged 6-16 years understood this norm, but that those aged 10-years and below seemed mainly to avoid expressing prejudice because they were aware that doing so would incur external disapproval. Contrastingly, those older than 10-years actually showed an implicit aversion to the concept of prejudice, consistent with the idea that they had genuinely internalised the anti-prejudice norm to the point where it influenced their reasoning (Rutland et al., 2005).

Much more work needs to be done in this sphere before any firm conclusions can be drawn, yet the suggestion that true normative internalisation does not occur until well into middle childhood is intriguing in light of an increasingly observed disconnect between younger children's cognitive *understanding* of social norms and their actual willingness to act upon them in circumstances where doing so negatively affects their own payoffs. Blake et al (2014) have referred to this dichotomy as the "knowledge-behaviour gap" and attribute it to two factors. Firstly, as discussed, young children appear to have a particularly strong sensitivity to relative evaluation, and a corresponding tendency to view outcomes in (objectively unnecessary) zero-sum terms. The second factor is that, compared to adults, children appear to become unusually competitive over resources which have been given to them by a third party, as compared with resources which they have worked together to acquire.

Both of these factors seem plausible motivators of competitive and hyper-competitive behaviour in young children. However, so does Haidt's suggestion that the individual does not develop the ability to actually *care* about social norms enough to train themselves into thinking and acting in accordance with them, until they each middle childhood. For example, middle childhood generally, and in particularly the age of 8-years specifically, has been repeatedly identified as a point in development at which a radical behavioural shift occurs away from selfish choices and towards those which correspond with norms of fairness and or/generosity (Blake & McAuliffe, 2011; Fehr et al., 2008; House et al., 2013; McAuliffe et al., 2013; McAuliffe et al., 2014; Sheskin, Bloom, et al., 2014; Steinbeis & Singer, 2013). Furthermore, a recent study (Smith et al., 2013) found that whilst all children across the age range of 3-8 years explicitly endorsed fair theoretical distributions of resources and predicted that others would do the same, only those at ages 7-8 were actually willing to share equally in practice. In other words, below the age of 7, children were unwilling to act in accordance with a simple norm they were perfectly capable of understanding, and even superficially endorsing.

Not all of this shift toward more normative behaviour in mid-to-late childhood is entirely due to a sudden surge in normative internalisation, of course. For instance, Shaw et al (2012) recently showed that a majority of all children in a sample aged 6-11 were equally prepared to make unfair distributions of resources to their own advantage, but that the elder children in the study were more likely to do so via a method that lent the prospect of plausible deniability to their selfishness. They were better at *appearing* normatively compliant, rather than genuinely more compliant. This does not represent true internalisation, which should lead to automatic compliance even in the absence of external compulsion (Gintis, 2007; Perugini et al., 2003; Rutland et al., 2005). Nevertheless, the weight of evidence in favour of a shift towards more normative behaviour beginning at around the age of 8-years is such that greater feelings of prosociality must play a part, and it is likely that internalisation of norms favouring pro-social behaviour contribute to these feelings.

8.2.4: Children's Attendance to Social Norms

There was less evidence for the impact of social norms upon children in Chapters 3-7 than there was for fairness concerns. This was not to say that such evidence was absent. There was evidence of a greater preference for making fair offers amongst older children in Experiment 1; for more punitive behaviour, both costly and non-costly in Experiments 1 and 2; and a reasonably high level of costly punishment Experiment 3. Most, if not all, of these behaviours were likely at least partially motivated by considerations of the "right" or "fair" thing to do, though it is true that fairness concerns will have also played a big part.

Generally, however, it would seem that children were often willing to discard normative behaviours of fair sharing and mutual cooperation in favour of competitive and hyper-competitive behaviours. Fairness concerns such as inequity aversion and relative evaluation appeared to play a greater role in their economic reasoning than social norms, certainly in the extent to which they caused violations of the SEM. This may have been because many of the subjects sampled, particularly those in early- to mid-childhood, had not yet undergone a full phase of normative internalisation. However, it may also be the case that norms are simply weaker motivators of other-regarding behaviour than are fairness concerns (Gächter, 2014). Previous attempts to compare the strengths of the two types of motivation experimentally have indeed reached a similar conclusion (Gächter, Nosenzo, & Sefton, 2013). This dissertation adds support to such suggestions by showing that certain fairness concerns, such as those concerning inequity and ranking, appear to be something to which young children

spontaneously attend to a greater degree than norms. We may be predisposed to attend to social norms in our pre-school years (Chudek & Henrich, 2011; Rakoczy & Schmidt, 2013), but their actual internalisation can take much longer than this to occur (Kogut, 2012; Rutland et al., 2005; Smith et al., 2013). Indeed, in some people this internalisation arguably fails to occur at all.

8.2.5: The Ultra-Social Animal?

A consistent picture is emerging from scrutiny of my results. High levels of theft; low levels of giving; high levels of punishment; low levels of positive reciprocity; an intolerance for negative inequity but a preference for advantageous inequity; zero-sum thinking; low levels of normative compliance; relative advantage seeking; the absence of social welfare concerns; conditional cooperation; a failure to establish long-term cooperative outcomes. Are these really the actions of an "ultra-social animal" (Hill, Barton, & Hurtado, 2009; Tomasello, 2014)?

The answer is yes. Whilst some of the behaviours witnessed in experimental Chapters 3-7, for example stealing in the MUGWT, are simply (strategically misguided) attempts to self-maximise in the manner the SEM would predict, others are much more complex behaviours unique to an unusually sociable animal. Punishment behaviours and relative evaluation in particularly are potentially costly actions which can and do frequently violate the SEM due specifically to a concern for others' welfare. Both such acts, and many others that humans routinely perform besides, are impossible without the presence of a target individual and, importantly, a concern with that individual's outcomes which is strong enough to overcome the desire to simply maximise one's own payoff. The important point is that these violations of the SEM occur in a manner that is hyper-competitive. As well as being the ultra-social animal, we are also unusually anti-social in our capacity for deriving satisfaction from others' losses, due to motivations such as envy and schadenfreude. A better label for *Homo sapiens*, therefore, might be "the other-regarding animal".

The purpose of this dissertation is not to deny that children have many prodigious prosocial and cooperative abilities. Even amongst pre-schoolers, examples are as varied as collaborating to achieve joint goals (Bullinger et al., 2011; Duguid et al., 2014), helping others achieve their desired goals (Warneken & Tomasello, 2006), displaying empathy towards the suffering of others (Svetlova et al., 2010; Vaish, Carpenter, & Tomasello, 2009), and plenty more besides. The extent to which we perform these social behaviours far exceeds that seen in our closest

evolutionary relatives (Tomasello, 2014; Tomasello, Melis, Tennie, Wyman, & Herrmann, 2012b). Rather than seeking to minimise these behaviours, my purpose is to highlight that on their own they are not necessarily sufficient to understand and explain cooperative and social behaviour. Instead, more attention needs to be paid to the factors that limit cooperation, and also whether and how it can survive in repeated interactions, as opposed to only isolated, stylised tasks (Bolton & Ockenfels, 2000).

8.2.6: Explaining Hyper-Competitiveness

The SEM was consistently violated across all eight of my experiments, in all cases due to the presence of negative other-regarding preferences. This begs an obvious question: why? When the SEM is violated for positive other-regarding reasons, we can point to a number of potential explanations, such as altruism (Batson & Shaw, 1991), direct reciprocity (Trivers, 1971), social norms (Perugini et al., 2003) and the presence and threat of social control/punishment (Fehr et al., 2002; Gächter et al., 2008)[12]. Many, if not all of these mechanisms are to some degree culturally learned. With negative other-regarding concerns, however, this is seems to be less consistently the case. Indeed, negative other-regarding concerns appear at first to be something of a puzzle. They are not economically rational but, unlike positive other-regarding concerns (which are also not economically rational), they are also often not always congruent with notions of normative behaviour, and therefore less likely to be deliberately taught to young children. Parents, teachers, childminders and other long-suffering agents of socialisation devote countless hours to encouraging and cajoling their charges into behaving hyper-cooperatively: to share, to empathise, to forgive, to be kind and to reciprocate kindness, to refrain from theft, bullying, violence and, generally, to play nicely. By contrast, whilst social norms generally dictate that social deviants should be punished, few children are encouraged to enact spiteful and mutually destructive punishment on their own initiative, and surely even fewer are encouraged to frame resource distribution in relative, zero-sum terms.

It seems curious, then, that the further back into childhood we go, the more we find higher levels of certain types of spiteful cognitions and preferences (Bügelmayer & Spiess, 2014; Häger, 2010; Sheskin, Bloom, et al., 2014; Steinbeis & Singer, 2013), much higher, indeed, than is the case in adults (Al-Ubaydli et al., 2013; Fehr et al., 2013). In the case of positive ORPs, the same is simply not true. Altruistic giving runs counter to self-interest and, consequently, takes

[12] And indeed several more which are beyond the immediate scope of this discussion, for example indirect reciprocity (Alexander, 1987) or between and within-groups conflict (Van Vugt, De Cremer, & Janssen, 2007).

a long time to properly mature (Benenson et al., 2007; Fehr et al., 2008; Fehr et al., 2013; House et al., 2013). However, behaviours such as sharing, helping and positive reciprocity are, on the whole, normatively encouraged by most if not all cultures (Haidt & Joseph, 2004). Spiteful income-reduction also runs counter to self-interest. However, its application is only considered normative in certain contexts, and yet it appears to be present to a greater extent from a younger age. So what is going on?

8.2.7: Two Types of Hyper-Competitiveness

The answer to this may lie in the fact that, as touched upon in Section 8.2.1 (p.182), violations of the SEM caused by negative ORPs appear to be split into two-types: normative, and non-normative. The first of these, normative violations, are not as problematic to explain. They are "normative" in that the spiteful act they require constitutes what is known as "altruistic" punishment (Fehr & Gächter, 2002). Punishment of this type is so-called because it contributes to maintenance of the social order by deterring free-riders and indeed, in a wider context beyond economics, deviants and criminals generally. Humans may be the only animal which engages in such acts (Riedl et al., 2012, 2015). Children may not be socialised into "spitefully punishing people" (or at least certainly not in such terms), but they do tend to be encouraged to adhere to society's rules and to help ensure that these rules are upheld by others. Altruistic punishment, like altruistic giving, violates immediate self-interest and may therefore take a while to be internalised in the growing mind. Experiments such as the moonlighting game in Chapter 5, however, suggest that this may happen earlier than is the case with altruistic giving. Nevertheless, punishment, like altruism needs to be learned.

This explanation, however, cannot explain the zero-sum thinking and relative advantage seeking reported in Chapters 6 and 7. This is not normative behaviour and, even if it was, many of the children in these experiments had not yet reached the age of norm internalisation anyway. So, if socialisation is not the answer, then why do we observe the high frequency of negative ORPs in so many young children in social and economic situations? What is the cause of "non-normative" negative ORPs? The eight experiments presented here are suggestive of the importance of one factor in particular. This is the tendency, demonstrated by Bolton and Ockenfels' (2000), to attend to "relative payoff standing", which appears to be peculiarly strong in humans, and probably has some innate basis. It seems extremely unlikely that the children as young as 4-5 years old who were repeatedly observed defecting in experiments such as the battle of the sexes game and stag hunt were doing so because they had been

extensively socialised into prioritising relative advantage over mutualistic cooperation. Instead, their tendency to evaluate the situation relatively and act accordingly appears to have been intuitive, much in the way that proponents of social comparison theory would suggest (Festinger, 1954; Fiske, 2011; Mussweiler, 2003).

The key thing which characterises human hyper-cooperativeness is the extent to which we are willing to cooperate with others to whom we are not genetically related (Hill et al., 2009). This allows for great benefits, but also opens us up to the risk of being exploited by free-riders. It may be for this reason that we are also exceptionally sensitive to how our outcomes compare to those of others. Indeed, our capacity to detect "cheats" is another prominent candidate for an innate cognitive ability (Cosmides & Tooby, 1992, 2008). A tendency to attend to who has more than us and why, relative evaluation in other words, would be an essential precursor of this ability, since in its absence it would be impossible to keep track of who is cheating and who is not. Targeted second and third party punishment, often "altruistic" in scope, is a key part of the human social order (Fehr & Gächter, 2002; Gächter, 2014). Being attuned to relative standing may be one of punishment's cognitive drivers for two reasons. The first is that is allows an accurate understanding of when a target has more than oneself. The second is that by flagging this as an aversive stimulus it can produce the motivation for a functionally spiteful response. Only through social learning of norms and cognitive maturation, however, do we learn to target this proclivity appropriately.

Of course, an ability to assess relative standing could have additional uses beyond simply motivating punishment. Comparing outcomes relatively can also serve a useful offensive function by telling us when it might be a good idea to engage in some free-riding of our own, at the expense of comparatively wealthy, genial and/or naïve targets. Furthermore, relative standing extends well beyond economics to more general social constructs such as hierarchy and influence. Acquiring both of these can be extremely useful predictors of future income and success, so perhaps it is little wonder that people appear to value them as ends in their own right. It may sometimes be worth losing an eye if, as part of the same arrangement, one can become ruler in the Kingdom of the Blind. Loewenstein et al (1989) referred to people whose thinking was categorised by this type of attitude, around a third of their sample, as "ruthless competitors". Certainly, a sub-set of the population continues to think this way well into adulthood (Abbink & Herrmann, 2011; Herrmann & Orzen, 2008), but it is an outlook which appears to be present amongst many more of us in our pre-adolescent years.

8.3: Methodological Considerations

Chapter 2 outlines the methodological considerations of this dissertation in some detail. Nevertheless there are some aspects which require further discussion in the context of the experimental results. The first is that the experiments were conducted in what can be termed a "semi-public" setting, meaning that subjects in the Live Science area were sometimes being watched, and were aware of this. I can think of no obvious reason to suppose that this arrangement would have artificially inflated the levels of hyper-competitiveness on display. The existing literature on the effect of observers upon adults' behaviours overwhelmingly concludes that it makes people *more* pro-social, not less (Andreoni & Petrie, 2004; Bateson, Nettle, & Roberts, 2006; Hardy & Van Vugt, 2006; Piazza & Bering, 2008; Satow, 1975). The existing research in the developmental literature presents a very similar picture in children aged 5-years and upwards (Engelmann et al., 2012; Engelmann et al., 2013; Leimgruber et al., 2012; A. Shaw et al., 2014). Consequently, it seems unlikely that the presence of observers caused the children in my sample to behave more competitively than they otherwise would have done.

A more plausible explanation for the perhaps unusually high levels of (hyper) competition on display is the way in which points were distributed. Paying children in economics experiments with a pseudo-currency with which they can later purchase prizes is common practice (Fan, 2000; Harbaugh & Krause, 2000; House et al., 2013; Leimgruber et al., 2012; Robbins & Rochat, 2011). More unusual, however, is having physical piles of this currency building up next to participants, in open view of all. Indeed, I know of no other existing study in which this has been done. There is a good chance that allowing overt relative comparison in this manner primed competitive motives in children given that they clearly attend very strongly to relative outcomes, although since there is no experimental evidence on this topic this is speculation.

Whilst acknowledging that this aspect of my methodology may have played its part in inflating hyper-competitiveness to levels above what might otherwise have been seen, I still feel it was justified. Firstly, I wanted to make absolutely sure that children had a meaningful guide as to how they were doing. Some of my experiments involved large numbers of points and it was felt that children might struggle to keep track of their scores, and what they actually meant, without a visual clue to provide context. "Filling the bucket" seemed a good, simple guide. Secondly, I would argue that rewarding children publically in this manner is more naturalistic

than the alternative. In their everyday lives, children are rarely paid by invisible deposits into bank accounts, or by shady hidden payments in brown envelopes. Instead, in contexts such as school assemblies and birthday parties, they are presented with rewards physically and publically, in front of their peers. Applying the same logic to economic experiments therefore seems justified, or at least certainly no harder to justify than using a mysterious reward which alternately appears and disappears from sight and mind, in an experiment in which one's alleged partner may or may not exist.

An important point to add here is that if children really do have strong moral and normative preferences against scenarios such as advantageous inequity then these ought to be robust enough to reliably emerge in a wide array of contexts, including those like the present dissertation which were not necessarily conducive to encouraging them. Instead, they often fail to recur even when the same experiment is repeated with different samples (e.g. Fehr et al., 2008; Fehr et al., 2013).

A final issue is the limited control I had over the demographics of my subjects. On the whole, I cannot see any evidence that this radically affected the results of most of my experiments. Gender and participants' level of acquaintance were included as potential confounds in all of the analyses presented here, but did not generally exert much obvious effect on the overall findings except for in one or two isolated cases which I have reported. Care was taken to match participants for age within-dyads as closely as possible, yet at times this was not always possible. This, however, is not necessarily a weakness, as a recent developmental study on the trust game shows. Younger children in the trust game are generally less trustworthy (Evans et al., 2013; Sutter & Kocher, 2007; van den Bos et al., 2012). Interestingly, older subjects seem to be aware of this, and those in late childhood are markedly less trusting when paired with a younger child than with a peer. In other words, the age of their partner can systematically affect the children's economic behaviour (Evans et al., 2013). This means that if children are constantly paired only with other subjects of their own age, all of their decisions will be systematically biased to reflect this. Only by pairing, for example, 8-year olds with 6-year olds, with 8-year olds *and* with 10-year olds is it possible to get a true sense of the 8-year olds' overall mean preferences, rather than just a sense of how they play strategically against peers specifically. Consequently, the fact that some of my dyads were perfectly matched for age whilst others were not can actually be considered something of a strength, albeit an incidental one rather than a product of my own design.

A valid caveat to add on the topic of demographics, however, is that my sample was disproportionately WEIRD, meaning that subjects originated from a country which is Western, Educated, Industrialised, Rich, and Democratic (Henrich, Heine, & Norenzayan, 2010). It is therefore not a given that the patterns of hyper-competitiveness on display here would apply equally to all children. If anything, given that Western cultures tend to stress individualism more than others (Berry, 2002), it may be the case that children raised in such an environment are more prone to spitefully motivated, relative comparisons than counterparts from other-cultures. At least one-cross cultural study has looked into costly punishment behaviour, for example, concluding that it does indeed occur earlier in children from a Western background (Robbins & Rochat, 2011). More work of this type therefore needs to be done before fully informed conclusions can be drawn regarding children's hyper-competitive behaviours. One thing I will add in mitigation, however, is that London is one of the world's most ethnically and culturally diverse cities, and the Science Museum is a free attraction which makes a concerted effort to attract visitors and school groups from disparate backgrounds. Consequently, my sample was less overwhelmingly Caucasian and middle-class than is the case in most developmental work conducted in the West, although doubtless the prevalence of both demographics was still disproportionately high.

8.4: Implications and Ideas for Future Research

My findings show that children are highly prone to spiteful violations of the SEM throughout the span of pre-adolescence. Consequently, although previous work stressing their hyper-cooperative nature has demonstrated the emergence of many sociable and cooperative abilities, it may over-estimate the tendency of children to act on them outside of specific experimental paradigms. As a general suggestion, then, future experiments on sociality and cooperation need to go beyond comparing a pro-social condition and a neutral alternative, and instead include the additional possibility for subjects to show spiteful preferences. The contrast between the trust game and the moonlighting game provides a useful example. The former asks whether or not people will positively reciprocate. The latter asks the same question, but also a much broader one as to how common and strong this tendency is when compared to negative reciprocity. By providing both additional data and a greater sense of context, this tells us a lot more about human sociality than does the trust game by itself. Generally, then, I would call for the greater inclusion of spiteful and hyper-competitive options in experiments on children's cooperative behaviour, and indeed in experiments on cooperative behaviour generally, as well as more iterated experiments such as those of Chapters 6 and 7, in

order to more stringently probe the strength and persistence of cooperative preferences. In addition, I also propose a number of specific studies, based on the findings presented in this dissertation. The exploratory nature of my work has led to some unexpected findings, and many others which are in need of replication, verification and extension. Here are some of my own ideas.

Firstly, more work needs to be done into children's preferences for mutualistic cooperation. In particular, an experiment in which pairs of subjects demonstrably and consistently prefer a mutualistic outcome over two alternative outcomes which are either a) individualistic *and* b) confer relative advantage, is necessary before it can be concluded that those in early to mid-childhood actually prefer mutualistic outcomes. This is important, since mutualistic cooperation is suggested by some as an innate tendency (McLoone & Smead, 2014), and even as one of the evolved cornerstones of all human cooperation (Tomasello et al., 2012a).

Secondly, the results of this dissertation identified inequity aversion and relative standing as the primary motivators of children's negative ORPs. Inequity aversion is reasonably well studied in childhood, but relative standing rather less so. An experiment which explored the latter in more depth would thus be worth pursuing. An example might be a study examining how much of a resource or prize children are willing to give up in order to propel themselves up some kind of arbitrary ranking system, whether a desire to do so is present in pre-schoolers, and how this changes over ontogeny.

Thirdly, it would be beneficial for the field as a whole to have more specific knowledge about how children's preferences are affected by being watched, both by their fellow subjects and by others. Although it seems very likely that audience effects would make children more generous (Engelmann et al., 2012; Engelmann et al., 2013; A. Shaw et al., 2014), there must be some point in development when a truly radical shift occurs here. Being made to make decisions publically consistently makes adults a lot more generous (Andreoni & Petrie, 2004; Bateson et al., 2006; Hardy & Van Vugt, 2006; Piazza & Bering, 2008; Satow, 1975). Though there is evidence that pre-adolescents are susceptible to this to an extent, the effect is nowhere near as pronounced, as this dissertation shows. When exactly does this begin to change? Pro-sociality experiments drawing samples from different phases of adolescence can help to answer this question.

This highlights another general area in which more work is needed. There are thousands of experiments on adults' economic preferences, and an increasing number on when they

emerge in young children, and how they develop throughout childhood. However, the period in-between remains comparatively unexplored, even though experiments which do compare teenagers to adults and/or children consistently show that their behaviour differs systemically from one or both of these groups (e.g. Fehr et al., 2013; Gummerum & Chu, 2014; Güroğlu et al., 2009; Peters et al., 2004; Sutter & Kocher, 2007; van den Bos et al., 2012). Experimental and neuroscientific work shows that the social brain continues to mature during the teenage years both physically and cognitively (Blakemore et al., 2007; Dumontheil et al., 2010; P. Shaw et al., 2008; Steinbeis et al., 2012; Steinbeis & Singer, 2013). Consequently, more work needs to be done here to look at other-regarding preferences which have not emerged by late childhood, such as outcome/intention differentiation, consistent positive reciprocity, and a concern for maximising social welfare.

Fourthly, children in this dissertation were reasonably consistent punishers, but they needed to be, since their peers consistently behaved in ways which warranted punishment. This raises an interesting question: does punishment actually work amongst children? If not, at what age does it start to do so? Is the same true of the threat of punishment? For example, theft behaviour increased in frequency over the course of repeated rounds in the moonlighting game, and was therefore not deterred by negative reciprocity. The literature on costly punishment in adults clearly shows that punishment is an effective stabiliser of cooperation (Chaudhuri, 2011; Fehr & Gächter, 2000; Gächter, 2014; Ledyard, 1997). Indeed, even the existence of punitive mechanisms, and the implied threat they create, is enough to hugely increase cooperation amongst adults (Gächter et al., 2008). The fact that in this developmental study they failed to do so is therefore intriguing, and suggests that a proper sensitivity to the threat of costly punishment does not seem to emerge until a later stage in development than a willingness to make punitive decisions. It should be noted that my finding here is contradicted by a recent study showing that the threat of punishment actually had quite a marked positive impact on cooperative behaviour amongst children aged between 7-11 years (Lergetporer et al., 2014). Nevertheless, whilst there is a growing literature on children's willingness to punish, there is less on its actual efficacy in terms of how it affects their behaviour in a strategic context. It would certainly be an interesting topic of future research to see whether punishment deters free-riding in children in the same way as it repeatedly does amongst adults, at what age it starts to do so, and how this tendency changes throughout development. Given children's increasingly well documented ability to cooperate conditionally, the public goods game seems a good candidate for the experiment to use here, since it would allow maximum comparability with existing data on adults. By doing this we could glean a sense of

how children's tendencies to react to spiteful violations of the SEM change in tandem with their willingness to enact them.

8.5: Conclusion

All living organisms are competitive, but perhaps only humans are hyper-competitive. This dissertation aimed to expand understanding of this phenomenon by focusing on the relatively unexplored area of its emergence and development in childhood. Strong evidence of hyper-competitiveness was in evidence across eight experimental contexts with children spread across a wide age range, spanning from pre-school to early adolescence. Indeed, hyper-competitive behaviours were more common than hyper-cooperative behaviours in the game theoretical contexts examined. Two types of hyper-competitive behaviour were observed. The first type, involving costly punishment, is possible to reconcile with existing conceptions of normative socialisation. However, the second type, involving zero-sum thinking and relative advantage seeking, is not. The key motivating factor behind these behaviours seemed to be an unusual, and certainly irrational, tendency to consider outcomes in a zero-sum manner, with high levels of relative social comparison and little concern for overall social welfare, or for cooperating with others for mutual gain. Pre-adolescence, it would seem, is a period of life characterised by high levels of negative other-regarding preferences and aggressive social comparisons. Normative socialisation seems to play little role in such behaviour. Consequently, whilst ultra-social, humans are also ultra-anti-social, and this is certainly true of our formative years. We are the other-regarding animal.

References

Abbink, K., & Herrmann, B. (2011). The moral costs of nastiness. *Economic Inquiry, 49*(2), 631-633.

Abbink, K., Irlenbusch, B., & Renner, E. (2000). The moonlighting game: An experimental study on reciprocity and retribution. *Journal of Economic Behavior & Organization, 42*(2), 265-277.

Al-Ubaydli, O., Jones, G., & Weel, J. (2013). Patience, cognitive skill, and coordination in the repeated stag hunt. *Journal of Neuroscience, Psychology, and Economics, 6*(2), 71-96.

Alencar, A. I., Siqueira, J. D., & Yamamoto, M. E. (2008). Does group size matter? Cheating and cooperation in Brazilian school children. *Evolution and Human Behavior, 29*(1), 42-48. doi: 10.1016/j.evolhumbehav.2007.09.001

Alexander, R. D. (1987). *The biology of moral systems*: Transaction Publishers.

Andreoni, J. (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *The Economic Journal, 100*(401), 464-477. doi: 10.2307/2234133

Andreoni, J. (1995). Cooperation in Public Goods Experiments: Kindness or Confusion? *American Economic Review, 85*(4), 891-904.

Andreoni, J., Harbaugh, W. T., & Vesterlund, L. (2002). The carrot or the stick: Rewards, punishments and cooperation. *University of Oregon Department of Economics Working Paper*(2002-1).

Andreoni, J., & Petrie, R. (2004). Public goods experiments without confidentiality: a glimpse into fund-raising. *Journal of public Economics, 88*(7), 1605-1623.

Angerer, S., Glätzle-Rützler, D., Lergetporer, P., & Sutter, M. (2015). Cooperation and discrimination within and across language borders: Evidence from children in a bilingual city: Ifo Institute for Economic Research at the University of Munich.

Archetti, M., & Scheuring, I. (2012). Review: Game theory of public goods in one-shot social dilemmas without assortment. *Journal of theoretical biology, 299*, 9-20.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *science, 211*(4489), 1390-1396.

Bardsley, N. (2008). Dictator game giving: altruism or artefact? *Experimental Economics, 11*(2), 122-133.

Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology letters, 2*(3), 412-414.

Batson, C. D., & Shaw, L. L. (1991). Evidence for altruism: Toward a pluralism of prosocial motives. *Psychological Inquiry, 2*(2), 107-122.

Baumard, N., André, J.-B., & Sperber, D. (2013). A mutualistic approach to morality: the evolution of fairness by partner choice. *Behavioral and brain sciences, 36*(01), 59-78.

Benenson, J. F., Pascoe, J., & Radmore, N. (2007). Children's altruistic behavior in the dictator game. *Evolution and Human Behavior, 28*(3), 168-175.

Bereby-Meyer, Y., & Fiks, S. (2013). Changes in negative reciprocity as a function of age. *Journal of Behavioral Decision Making, 26*(4), 397-403.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior, 10*(1), 122-142.

Berry, J. W. (2002). *Cross-cultural psychology: Research and applications*: Cambridge University Press.

Birch, L. L., & Billman, J. (1986). Preschool children's food sharing with friends and acquaintances. *Child development, 57*(2), 387-395. doi: doi: 10.2307/1130594

Blake, P. R., & McAuliffe, K. (2011). “I had so much it didn’t seem fair”: Eight-year-olds reject two forms of inequity. *Cognition, 120*(2), 215-224.

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.

Blakemore, S.-J., den Ouden, H., Choudhury, S., & Frith, C. (2007). Adolescent development of the neural circuitry for thinking about intentions. *Social cognitive and affective neuroscience, 2*(2), 130-139.

Boehm, C. (2012). *Moral origins: The evolution of virtue, altruism, and shame*. New York: Basic Books.

Bogin, B. (1990). The evolution of human childhood. *Bioscience, 40*(1), 16-25.

Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review, 90*(1), 166-193.

Bornstein, G., Mingelgrin, D., & Rutte, C. (1996). The effects of within-group communication on group decision and individual choice in the assurance and chicken team games. *Journal of Conflict Resolution, 40*(3), 486-501.

Bosman, R., Sutter, M., & van Winden, F. (2005). The impact of real effort and emotions in the power-to-take game. *Journal of economic psychology, 26*(3), 407-429.

Bosman, R., & Van Winden, F. (2000). The behavioral impact of emotions in a power-to-take game: An experimental study.

Bosman, R., & Van Winden, F. (2002). Emotional Hazard in a Power to Take Experiment. *The economic journal, 112*(476), 147-169.

Boyd, R., & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and sociobiology, 13*(3), 171-195.

Boyer, T. W. (2006). The development of risk-taking: A multi-perspective review. *Developmental Review, 26*(3), 291-345.

Brosnan, S. F., & De Waal, F. B. (2003). Monkeys reject unequal pay. *Nature*,*425*(6955), 297-299.

Brosnan, S. F., Parrish, A., Beran, M. J., Flemming, T., Heimbauer, L., Talbot, C. F., Wilson, B. J. (2011). Responses to the Assurance game in monkeys, apes, and humans using equivalent procedures. *Proceedings of the National Academy of Sciences, 108*(8), 3442-3447.

Bügelmayer, E., & Spiess, C. K. (2014). Spite and cognitive skills in preschoolers. *Journal of economic psychology, 45*, 154-167.

Bullinger, A. F., Wyman, E., Melis, A. P., & Tomasello, M. (2011). Coordination of chimpanzees (Pan troglodytes) in a stag hunt game. *International Journal of Primatology, 32*(6), 1296-1310.

Burger, J. M., Sanchez, J., Imberi, J. E., & Grande, L. R. (2009). The norm of reciprocity as an internalized social norm: Returning favors even when no one finds out. *Social Influence, 4*(1), 11-17.

Burkart, J., Allon, O., Amici, F., Fichtel, C., Finkenwirth, C., Heschl, A., Martins, E. (2014). The evolutionary origin of human hyper-cooperation. *Nature communications, 5*.

Buttelmann, D., & Böhm, R. (2014). The ontogeny of the motivation that underlies in-group bias. *Psychological science, 25*(4), 921-927.

Buttelmann, D., Zmyj, N., Daum, M., & Carpenter, M. (2013). Selective imitation of in-group over out-group members in 14-month-old infants. *Child development, 84*(2), 422-428.

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in cognitive sciences, 12*(5), 187-192.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. New Jersey: Princeton University Press.

Camerer, C., & Fehr, E. (2006). When does" economic man" dominate social behavior? *science, 311*(5757), 47-52.

Castelli, I., Massaro, D., Bicchieri, C., Chavez, A., & Marchetti, A. (2014). Fairness Norms and Theory of Mind in an Ultimatum Game: Judgments, Offers, and Decisions in School-Aged Children. *PLoS One, 9*(8), e105024. . doi: 10.1371/journal.pone.0105024

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics, 117*(3), 817-869.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics, 14*(1), 47-83.

Chen, X., Sasaki, T., Brännström, Å., & Dieckmann, U. (2015). First carrot, then stick: how the adaptive hybridization of incentives promotes cooperation. *Journal of The Royal Society Interface, 12*(102), 20140935.

Chudek, M., & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in cognitive sciences, 15*(5), 218-226.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides & J. Tooby (Eds.), *The adapted mind. Evolutionary psychology and the generation of culture.* (pp. 163-228). New York: Oxford University Press.

Cosmides, L., & Tooby, J. (2008). Can a general deontic logic capture the facts of human moral reasoning? How the mind interprets social exchange rules and detects cheaters. In W. Sinnott-Armstrong (Ed.), *The Evolution of Morality: Adaptations and Innateness* (Vol. 1, pp. 53-120). Cambridge, Massachusetts: MIT Press.

Brosnan, S. F., & De Waal, F. B. (2003). Monkeys reject unequal pay. *Nature*,*425*(6955), 297-299.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2008). Representative trust and reciprocity: prevalence and determinants. *Economic Inquiry, 46*(1), 84-90.

Dreber, A., Fudenberg, D., & Rand, D. G. (2014). Who cooperates in repeated games: The role of altruism, inequity aversion, and demographics. *Journal of Economic Behavior & Organization, 98*, 41-55.

Duguid, S., Wyman, E., Bullinger, A. F., Herfurth-Majstorovic, K., & Tomasello, M. (2014). Coordination strategies of chimpanzees and human children in a Stag Hunt game. *Proceedings of the Royal Society B-Biological Sciences, 281*(1796).

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.

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (1998). Prosocial development. Social, Emotional and Personality Development In W. Damon & R. R. Lerner (Eds.), *Handbook of child psychology*. Chichester: John Wiley and Sons.

Engel, C. (2011). Dictator games: a meta study. *Experimental Economics, 14*(4), 583-610.

Engelmann, J. M., Herrmann, E., & Tomasello, M. (2012). Five-year olds, but not chimpanzees, attempt to manage their reputations. *PLoS One, 7*(10), e48433.

Engelmann, J. M., Over, H., Herrmann, E., & Tomasello, M. (2013). Young children care more about their reputation with ingroup members and potential reciprocators. *Developmental science, 16*(6), 952-958.

Evans, A. M., Athenstaedt, U., & Krueger, J. I. (2013). The development of trust and altruism during childhood. *Journal of economic psychology, 36*, 82-95.

Falk, A., Fehr, E., & Fischbacher, U. (2003). On the nature of fair behavior. *Economic Inquiry, 41*(1), 20-26.

Falk, A., Fehr, E., & Fischbacher, U. (2008). Testing theories of fairness—Intentions matter. *Games and Economic Behavior, 62*(1), 287-303.

Fan, C.-P. (2000). Teaching children cooperation—An application of experimental game theory. *Journal of Economic Behavior & Organization, 41*(3), 191-209.

Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature, 454*(7208), 1079-1083.

Fehr, E., & Fischbacher, U. (2002). Why social preferences matter–the impact of non-selfish motives on competition, cooperation and incentives. *The economic journal, 112*(478), C1-C33.

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature, 425*(6960), 785-791.

Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human nature, 13*(1), 1-25.

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments.

American Economic Review, 90(4), 980-994.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature, 415*(6868), 137-140.

Fehr, E., Glätzle-Rützler, D., & Sutter, M. (2013). The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence. *European Economic Review, 64*, 369-383.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly journal of Economics, 114*(3), 817-868.

Festinger, L. (1954). A theory of social comparison processes. *Human relations, 7*(2), 117-140.

Field, A. (2009). *Discovering statistics using SPSS* (Third ed.). Los Angeles: Sage publications.

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters, 71*(3), 397-404.

Fiske, S. T. (2011). *Envy up, scorn down: How status divides us*. New York: Russell Sage Foundation.

Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior, 6*(3), 347-369.

Frank, R. H. (1987). If homo economicus could choose his own utility function, would he want one with a conscience? *The American Economic Review, 77*(4), 593-604.

Gächter, S., Herrmann, B., & Thöni, C. (2010). Culture and cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365*(1553), 2651-2661.

Gächter, S., Nosenzo, D., & Sefton, M. (2013). Peer effects in prosocial behaviour: social norms or social preferences? *Journal of the European Economic Association, 11*(3), 548-573.

Gächter, S., Renner, E., & Sefton, M. (2008). The long-run benefits of punishment. *science, 322*(5907), 1510-1510.

Gächter, S. (2014). Human Pro-Social Motivation and the Maintenance of Social Order. In E. Zamir & D. Teichman (Eds.), *Handbook on Behavioral Economics and the Law* (pp. 28-60). Oxford: Oxford University Press.

Gigerenzer, G. (2008). *Gut feelings: Short cuts to better decision making*. Harmondsworth: Penguin UK.

Gintis, H. (2000a). Beyond Homo economicus: evidence from experimental economics. *Ecological economics, 35*(3), 311-322.

Gintis, H. (2000b). Strong reciprocity and human sociality. *Journal of theoretical biology, 206*(2), 169-179.

Gintis, H. (2007). Unifying the behavioral sciences II. *Behavioral and brain sciences, 30*(01), 45-53.

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior, 24*(3), 153-172.

Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American sociological review, 25*, 161-178.

Grueneisen, S., Wyman, E., & Tomasello, M. (2015). Children use salience to solve coordination problems. *Developmental science, 18*(3), 495-501.

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.

Gummerum, M., Hanoch, Y., Keller, M., Parsons, K., & Hummel, A. (2010). Preschoolers' allocations in the dictator game: The role of moral emotions. *Journal of economic psychology, 31*(1), 25-34.

Gummerum, M., Takezawa, M., & Keller, M. (2009). The influence of social category and reciprocity on adults' and children's altruistic behavior. *Evolutionary Psychology, 7*(2), 295-316.

Gürerk, Ö., Irlenbusch, B., & Rockenbach, B. (2006). The competitive advantage of sanctioning institutions. *science, 312*(5770), 108-111.

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.

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

Häger, K. (2010). Envy and altruism in children *Jena economic research papers*. Jena: Freidrich Schiller University

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological review, 108*(4), 814.

Haidt, J. (2004). The Emotional Dog Gets Mistaken for a Possum. *Review of General Psychology, 8*(4), 283-290.

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus, 133*(4), 55-66.

Harbaugh, W. T., & Krause, K. (2000). Children's altruism in public good and dictator experiments. *Economic Inquiry, 38*(1), 95-109.

Harbaugh, W. T., Krause, K., & Liday, S. J. (2003). Bargaining by children. *University of Oregon Economics Working Paper*(2002-4).

Hardy, C. L., & Van Vugt, M. (2006). Nice guys finish first: The competitive altruism hypothesis. *Personality and Social Psychology Bulletin, 32*(10), 1402-1413.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review, 91*(2), 73-78.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., . . . Ensminger, J. (2005). "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and brain sciences, 28*(06), 795-815.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and brain sciences, 33*(2-3), 61-83.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., . . . Henrich, N. (2006). Costly punishment across human societies. *science, 312*(5781), 1767-1770.

Herrmann, B., & Orzen, H. (2008). The appearance of homo rivalis: Social preferences and the nature of rent seeking *CeDEx discussion paper series* (Vol. No 2008-10). Nottingham: University of Nottingham.

Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *science, 319*(5868), 1362-1367.

Hertwig, R., & Herzog, S. M. (2009). Fast and frugal heuristics: Tools of social rationality. *Social Cognition, 27*(5), 661-698.

Hill, K., Barton, M., & Hurtado, A. M. (2009). The emergence of human uniqueness: characters underlying behavioral modernity. *Evolutionary Anthropology, 18*(5), 187-200.

House, B., Henrich, J., Sarnecka, B., & Silk, J. B. (2013). The development of contingent reciprocity in children. *Evolution and Human Behavior, 34*(2), 86-93.

House, B., Silk, J. B., Henrich, J., Barrett, H. C., Scelza, B. A., Boyette, A. H., Laurence, S. (2013). Ontogeny of prosocial behavior across diverse societies. *Proceedings of the National Academy of Sciences, 110*(36), 14586-14591.

Jacobs, J. E., & Klaczynski, P. A. (2002). The development of judgment and decision making during childhood and adolescence. *Current directions in psychological science, 11*(4), 145-149.

Jensen, K. (2010). Punishment and spite, the dark side of cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365*(1553), 2635-2650.

Jensen, K., Call, J., & Tomasello, M. (2007a). Chimpanzees are rational maximizers in an ultimatum game. *science, 318*(5847), 107-109.

Jensen, K., Call, J., & Tomasello, M. (2007b). Chimpanzees are vengeful but not spiteful. *Proceedings of the National Academy of Sciences, 104*(32), 13046-13050.

Jensen, K., Hare, B., Call, J., & Tomasello, M. (2006). What's in it for me? Self-regard precludes altruism and spite in chimpanzees. *Proceedings of the Royal Society of London B: Biological Sciences, 273*(1589), 1013-1021.

Jensen, K., Vaish, A., & Schmidt, M. F. (2014). The emergence of human prosociality: aligning with others through feelings, concerns, and norms. *Frontiers in psychology, 5*.

Johnston, A. M., Johnson, S. G., Koven, M. L., & Keil, F. C. (2015). *Probabilistic Versus Heuristic Accounts of Explanation in Children: Evidence from a Latent Scope Bias.* Paper presented at the Proceedings of the 37th Annual Conference of the Cognitive Science Society.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society, 47*(2), 263-291.

Kaiser, I., Jensen, K., Call, J., & Tomasello, M. (2012). Theft in an ultimatum game: chimpanzees and bonobos are insensitive to unfairness. *Biology letters, 8*(6), 942-945.

Kanngiesser, P., Rossano, F., & Tomasello, M. (2015). Late Emergence of the First Possession Heuristic: Evidence From a Small-Scale Culture. *Child development, 86*(4), 1282-1289.

Keil, L. J. (1986). Rules, reciprocity, and rewards: A developmental study of resource allocation in social interaction. *Journal of Experimental Social Psychology, 22*(5), 419-435.

Kenward, B., & Östh, T. (2012). Enactment of third-party punishment by 4-year-olds. *Frontiers in psychology, 3*(373), 1-9.

Keysar, B., Converse, B. A., Wang, J., & Epley, N. (2008). Reciprocity Is Not Give and Take Asymmetric Reciprocity to Positive and Negative Acts. *Psychological science, 19*(12), 1280-1286.

Kogut, T. (2012). Knowing what I should, doing what I want: From selfishness to inequity aversion in young children's sharing behavior. *Journal of economic psychology, 33*(1), 226-236.

Krueger, J. I., Massey, A. L., & DiDonato, T. E. (2008). A matter of trust: From social preferences to the strategic adherence to social norms. *Negotiation and Conflict Management Research, 1*(1), 31-52.

Kurzban, R., Burton-Chellew, M. N., & West, S. A. (2015). The Evolution of Altruism in Humans. *Annual review of psychology, 66*, 575-599.

Kuzawa, C. W., Chugani, H. T., Grossman, L. I., Lipovich, L., Muzik, O., Hof, P. R., . . . Lange, N. (2014). Metabolic costs and evolutionary implications of human brain development. *Proceedings of the National Academy of Sciences, 111*(36), 13010-13015.

Lagattuta, K. H. (2014). Linking past, present, and future: Children's ability to connect mental states and emotions across time. *Child Development Perspectives, 8*(2), 90-95.

Lamba, S., & Mace, R. (2010). People recognise when they are really anonymous in an economic game. *Evolution and Human Behavior, 31*(4), 271-278.

Ledyard, J. (1997). Public goods: A survey of experimental research: David K. Levine.

Leimgruber, K. L., Shaw, A., Santos, L. R., & Olson, K. R. (2012). Young children are more generous when others are aware of their actions. *PLoS One, 7*(10), e48292.

Lergetporer, P., Angerer, S., Glätzle-Rützler, D., & Sutter, M. (2014). Third-party punishment increases cooperation in children through (misaligned) expectations and conditional cooperation. *Proceedings of the National Academy of Sciences, 111*(19), 6916-6921.

Levine, D. K. (1998). Modeling altruism and spitefulness in experiments. *Review of economic dynamics, 1*(3), 593-622.

Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *The journal of economic perspectives, 21*(2), 153-174.

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.

Loewenstein, G. F., Thompson, L., & Bazerman, M. H. (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social psychology, 57*(3), 426-441.

Machiavelli, N. (1950). *The Prince, and The Discourses*. New York: Random House.

McAuliffe, K., Blake, P. R., Kim, G., Wrangham, R. W., & Warneken, F. (2013). Social influences on inequity aversion in children. *PLoS One, 8*(12), e80966.

McAuliffe, K., Blake, P. R., & Warneken, F. (2014). Children reject inequity out of spite. *Biology letters, 10*(12), 20140743.

McAuliffe, K., Jordan, J. J., & Warneken, F. (2015). Costly third-party punishment in young children. *Cognition, 134*, 1-10.

McCrink, K., Bloom, P., & Santos, L. R. (2010). Children's and adults' judgments of equitable resource distributions. *Developmental Science, 13*(1), 37-45.

McLoone, B., & Smead, R. (2014). The ontogeny and evolution of human collaboration. *Biology & Philosophy, 29*(4), 559-576.

Melis, A. P., Altrichter, K., & Tomasello, M. (2013). Allocation of resources to collaborators and free-riders in 3-year-olds. *Journal of experimental child psychology, 114*(2), 364-370.

Moore, A. M., & Ashcraft, M. H. (2015). Children's mathematical performance: Five cognitive tasks across five grades. *Journal of experimental child psychology, 135*, 1-24.

Moore, C. (2009). Fairness in children's resource allocation depends on the recipient. *Psychological science, 20*(8), 944-948.

Mueller. (2003). *Public choice III*. Cambridge: Cambridge University Press.

Murnighan, J. K., & Saxon, M. S. (1998). Ultimatum bargaining by children and adults. *Journal of economic psychology, 19*(4), 415-445.

Mussweiler, T. (2003). Comparison processes in social judgment: mechanisms and consequences. *Psychological review, 110*(3), 472.

Nelson, W. R. (2002). Equity or intention: it is the thought that counts. *Journal of Economic Behavior & Organization, 48*(4), 423-430.

Neugebauer, T., Perote, J., Schmidt, U., & Loos, M. (2009). Selfish-biased conditional cooperation: On the decline of contributions in repeated public goods experiments. *Journal of economic psychology, 30*(1), 52-60.

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *science, 314*(5805), 1560-1563.

Olson, K. R., & Spelke, E. S. (2008). Foundations of cooperation in young children. *Cognition, 108*(1), 222-231.

Ortony, A., Clore, G. L., & Collins, A. (1990). *The cognitive structure of emotions*. Cambridge: Cambridge University Press.

Perugini, M., Gallucci, M., Presaghi, F., & Ercolani, A. P. (2003). The personal norm of reciprocity. *European Journal of Personality, 17*(4), 251-283.

Peters, H. E., Ünür, A. S., Clark, J., & Schulze, W. D. (2004). Free-Riding and the Provision of Public Goods in the Family: A Laboratory Experiment*. *International Economic Review, 45*(1), 283-299.

Peysakhovich, A., Nowak, M. A., & Rand, D. G. (2014). Humans display a 'cooperative phenotype'that is domain general and temporally stable. *Nature communications, 5*(4939).

Piazza, J., & Bering, J. M. (2008). Concerns about reputation via gossip promote generous allocations in an economic game. *Evolution and Human Behavior, 29*(3), 172-178.

Price, S. A., & Brosnan, S. F. (2012). To each according to his need? Variability in the responses to inequity in non-human primates. *Social Justice Research*, *25*(2), 140-169.

Proctor, D., Williamson, R. A., de Waal, F. B., & Brosnan, S. F. (2013). Chimpanzees play the ultimatum game. *Proceedings of the National Academy of Sciences, 110*(6), 2070-2075.

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.

Raihani, N., & McAuliffe, K. (2012). Human punishment is motivated by inequity aversion, not a desire for reciprocity. *Biology letters, 8*(5), 802-804.

Rakoczy, H., & Schmidt, M. F. (2013). The early ontogeny of social norms. *Child Development Perspectives, 7*(1), 17-21.

Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: young children's awareness of the normative structure of games. *Developmental psychology, 44*(3), 875.

Rege, M., & Telle, K. (2004). The impact of social approval and framing on cooperation in public good situations. *Journal of public Economics, 88*(7), 1625-1644.

Reyna, V. F. (2012). A new intuitionism: Meaning, memory, and development in Fuzzy-Trace Theory. *Judgment and Decision Making, 7*(3), 332- 359.

Reyna, V. F., & Ellis, S. C. (1994). Fuzzy-trace theory and framing effects in children's risky decision making. *Psychological science, 5*(5), 275-279.

Riedl, K., Jensen, K., Call, J., & Tomasello, M. (2012). No third-party punishment in chimpanzees. *Proceedings of the National Academy of Sciences, 109*(37), 14824-14829.

Riedl, K., Jensen, K., Call, J., & Tomasello, M. (2015). Restorative Justice in Children. *Current Biology, 25*(13), 1731–1735.

Robbins, E., & Rochat, P. (2011). Emerging signs of strong reciprocity in human ontogeny. *Frontiers in Psychology, 2*.

Rosati, A. G., Wobber, V., Hughes, K., & Santos, L. R. (2014). Comparative developmental psychology: How is human cognitive development unique? *Evolutionary Psychology, 12*(2), 448-473.

Rossano, F., Rakoczy, H., & Tomasello, M. (2011). Young children's understanding of violations of property rights. *Cognition, 121*(2), 219-227.

Rutland, A., Cameron, L., Milne, A., & McGeorge, P. (2005). Social Norms and Self-Presentation: Children's Implicit and Explicit Intergroup Attitudes. *Child development, 76*(2), 451-466.

Salali, G. D., Juda, M., & Henrich, J. (2015). Transmission and development of costly punishment in children. *Evolution and Human Behavior, 36*(2), 86-94.

Sally, D., & Hill, E. (2006). The development of interpersonal strategy: Autism, theory-of-mind, cooperation and fairness. *Journal of economic psychology, 27*(1), 73-97.

Sasaki, T., Brännström, Å., Dieckmann, U., & Sigmund, K. (2012). The take-it-or-leave-it option allows small penalties to overcome social dilemmas. *Proceedings of the National Academy of Sciences, 109*(4), 1165-1169.

Satow, K. L. (1975). Social approval and helping. *Journal of Experimental Social Psychology, 11*(6), 501-509.

Schmidt, M. F., & Sommerville, J. A. (2011b). Fairness expectations and altruistic sharing in 15-month-old human infants. *PLoS One, 6*(10), e23223. doi: 10.1371/journal.pone.0023223

Sebastián-Enesco, C., Hernández-Lloreda, M. V., & Colmenares, F. (2013). Two and a half-year-old children are prosocial even when their partners are not. *Journal of experimental child psychology, 116*(2), 186-198.

Sebastián-Enesco, C., & Warneken, F. (2015). The shadow of the future: 5-Year-olds, but not 3-year-olds, adjust their sharing in anticipation of reciprocation. *Journal of experimental child psychology, 129*, 40-54.

Shaw, A., Montinari, N., Piovesan, M., Olson, K. R., Gino, F., & Norton, M. I. (2014). Children develop a veil of fairness. *Journal of Experimental Psychology: General, 143*(1), 363.

Shaw, P., Kabani, N. J., Lerch, J. P., Eckstrand, K., Lenroot, R., Gogtay, N., Rapoport, J. L. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *The Journal of Neuroscience, 28*(14), 3586-3594.

Sheskin, M., Ashayeri, K., Skerry, A., & Santos, L. R. (2014). Capuchin monkeys (Cebus apella) fail to show inequality aversion in a no-cost situation. *Evolution and Human Behavior*, *35*(2), 80-88.

Sheskin, M., Bloom, P., & Wynn, K. (2014). Anti-equality: Social comparison in young children. *Cognition, 130*(2), 152-156.

Sheskin, M., Chevallier, C., Lambert, S., & Baumard, N. (2014). Life-history theory explains childhood moral development. *Trends in cognitive sciences, 18*(12), 613-615.

Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge: Cambridge University Press.

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

Smith, C. E., Blake, P. R., & Harris, P. L. (2013). I should but I won't: why young children endorse norms of fair sharing but do not follow them. *PLoS One, 8*(3), e59510.

Steelandt, S., Dufour, V., Broihanne, M.-H., & Thierry, B. (2012). Children Base Their Investment on Calculated Pay-Off. *PLoS One, 7*(3), e33239.

Steinbeis, N., Bernhardt, B. C., & Singer, T. (2012). Impulse control and underlying functions of the left DLPFC mediate age-related and age-independent individual differences in strategic social behavior. *Neuron, 73*(5), 1040-1051.

Steinbeis, N., & Singer, T. (2013). The effects of social comparison on social emotions and behavior during childhood: the ontogeny of envy and Schadenfreude predicts developmental changes in equity-related decisions. *Journal of experimental child psychology, 115*(1), 198-209.

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.

Sutter, M., & Kocher, M. G. (2007). Trust and trustworthiness across different age groups. *Games and Economic Behavior, 59*(2), 364-382.

Svetlova, M., Nichols, S. R., & Brownell, C. A. (2010). Toddlers' prosocial behavior: From instrumental to empathic to altruistic helping. *Child development, 81*(6), 1814-1827.

Sylwester, K., Herrmann, B., & Bryson, J. J. (2013). Homo homini lupus? Explaining antisocial punishment. *Journal of Neuroscience, Psychology, and Economics, 6*(3), 167-188.

Takagishi, H., Kameshima, S., Schug, J., Koizumi, M., & Yamagishi, T. (2010). Theory of mind enhances preference for fairness. *Journal of experimental child psychology, 105*(1), 130-137.

Takagishi, H., Koizumi, M., Fujii, T., Schug, J., Kameshima, S., & Yamagishi, T. (2014). The role of cognitive and emotional perspective taking in economic decision making in the ultimatum game. *PLoS One, 9*(9), e108462.

Tomasello, M. (2009). *Why we cooperate*. Cambridge, Massachussets: MIT press.

Tomasello, M. (2014). The ultra-social animal. *European journal of social psychology, 44*(3), 187-194.

Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012a). Two key steps in the evolution of human cooperation: the independece hypothesis. *Current anthropology, 53*(6), 673-692.

Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012b). Two key steps in the evolution of human cooperation: the independence hypothesis. *Current Anthropology, 53*(6), 673-692.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly review of biology, 46*(1), 35-57.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of business, 59*(4), S251-S278.

Vaish, A., Carpenter, M., & Tomasello, M. (2009). Sympathy through affective perspective taking and its relation to prosocial behavior in toddlers. *Developmental psychology, 45*(2), 534.

van den Bos, W., van Dijk, E., & Crone, E. A. (2012). Learning whom to trust in repeated social interactions: a developmental perspective. *Group Processes & Intergroup Relations, 15*(2), 243-256.

Van Vugt, M., De Cremer, D., & Janssen, D. P. (2007). Gender differences in cooperation and competition the Male-Warrior hypothesis. *Psychological science, 18*(1), 19-23.

Vogelsang, M., Jensen, K., Kirschner, S., Tennie, C., & Tomasello, M. (2014). Preschoolers are sensitive to free riding in a public goods game. *Frontiers in Psychology, 5*.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *science, 311*(5765), 1301-1303.

Warneken, F., & Tomasello, M. (2014). Extrinsic rewards undermine altruistic tendencies in 20-month-olds. *Developmental Psychology, 44*(6), 1785-1788.

Winking, J., & Mizer, N. (2013). Natural-field dictator game shows no altruistic giving. *Evolution and Human Behavior, 34*(4), 288-293.

Wischniewski, J., Windmann, S., Juckel, G., & Brüne, M. (2009). Rules of social exchange: game theory, individual differences and psychopathology. *Neuroscience & Biobehavioral Reviews, 33*(3), 305-313.

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.

Wyman, E., Rakoczy, H., & Tomasello, M. (2013). Non-verbal communication enables children's coordination in a "Stag Hunt" game. *European Journal of Developmental Psychology, 10*(5), 597-610.

Yamagishi, T., Horita, Y., Mifune, N., Hashimoto, H., Li, Y., Shinada, M., Simunovic, D. (2012). Rejection of unfair offers in the ultimatum game is no evidence of strong reciprocity. *Proceedings of the National Academy of Sciences, 109*(50), 20364-20368.

Zizzo, D. J., & Oswald, A. J. (2001). Are people willing to pay to reduce others' incomes? *Annales d'Economie et de Statistique, 63/64*, 39-65.