

How Experiments with Children Inform Economics

John A. List^a, Ragan Petrie^b, and Anya Samek^c

^aUniversity of Chicago and NBER

^bTexas A&M University

^cUniversity of California San Diego and NBER

February 2021

Abstract

In the past several decades the experimental method has lent deep insights into economics. One surprising area that has contributed is the experimental study of children, where advances as varied as the evolution of human behaviors that shape markets and institutions to how early life influences shape later life outcomes have been explored. We first develop a framework for economic preference measurement that provides a lens into how to interpret data from experiments with children. Next, we survey work that provides general empirical insights within our framework. Finally, we provide 10 tips for pulling off experiments with children, including factors such as taking into account child competencies, causal identification, and logistical issues related to recruitment and implementation. We envision the experimental study of children as a high growth research area in the coming decades as social scientists begin to more fully appreciate that children are active participants in markets who (might) respond predictably to economic incentives.

Keywords: children, economic experiments, economic preferences

JEL Codes: C9, D1, J1

Acknowledgements: Corresponding author: John List, University of Chicago, jlist@uchicago.edu. For helpful comments, we thank Steven Durlauf, Isabelle Brocas, Juan Carillo, Asadul Islam, Justin Holz, Udit Karna, Ariel Listo, participants at a variety of seminars/conference presentations, and especially three anonymous referees at the *Journal of Economic Literature*. For research assistance, we thank Kevin Sokal, Tarush Gupta, Andre Gray, Jessica Lopez, Andrew Yu, Haruka Uchida, Michael Cuna, Shreemayi Samujjwala and students at the Behavioral and Experimental Economics (BEE) research group. This research was funded by the Kenneth and Anne Griffin Foundation and by NIH grant 1R01DK114238.

1. Introduction

When spanning the research landscape over the past several decades it is not difficult to find examples of experimental studies where children play a relevant, and in many cases instrumental, role in the advancement of science. Many neuroscientists need children as experimental subjects. Developmental and social psychologists rely heavily on children too. Other scientists, from areas as varied as speech to autism to general medicine, critically depend on children to test their ideas. Indeed, in 1953 Jonas Salk injected his own three sons with a test polio vaccine based on an inactivate poliovirus. This early test paved the way for a nationwide trial that included nearly 2 million children. By 1955, the vaccine was widely reported safe and effective. While scientists across broad fields of study rely on children to gather data via experimentation, one might wonder—should economists conduct experiments with children too? On several dimensions the answer is unequivocally yes. To organize our thoughts, we classify why economists need children into two silos.

First, in the spirit of medical trials, economists, as well as neurologists, developmental psychologists, sociologists, and the like, must understand children themselves for immediate purposes. One prime example is experimental explorations in the economics of education wherein different elements in the environment are perturbed and outcomes measured (see, e.g., Sadoff, 2014 and Lavecchia et al., 2016 for excellent overviews).¹ Indeed, understanding models of human capital formation and early childhood education is an example of understanding children themselves and how best to invest in child development and skill formation. This is because children's choices are modeled as inputs into education production functions that shape their life outcomes (Heckman, 2006; List et al., 2018). Likewise, how to affect children's health and mental well-being are critical inputs into not only short-run outcomes but also those in the longer run. More generally, any interventions designed to inform policy prescriptions for children's well-being today fall into this first silo.

¹ The economics literature in this area is enormous, and spans experimental work using data from the seminal Perry Preschool and Abecedarian experiments (see Heckman et al., 2010; 2013; summaries are provided in Currie, 2001, Duncan and Magnuson, 2013 and Almond et al., 2018) to work in education research in the 1960s where some children were labeled as high performers and others low performers, when they had actually performed identically on achievement tests (Rosenthal and Jacobsen, 1968) to work evaluating educational interventions in a developing country context (see Kremer, 2003 for a summary).

Our second silo includes experiments exploring the behaviors and preferences of children, which we argue provide a unique glimpse into understanding adults, who are the primary decision-makers in most important markets. Much like the work in the medical sciences on normal paths of physiological maturation and determining the roots of illness and later life advantages and disadvantages, or the social science work on learned behaviors and language/speech development, economic experiments with children can focus on the decisions and preferences of children that potentially foreshadow adult behavior. As such, experiments with children can be thought of as complementary to recent approaches to understanding underlying factors shaping behavior, such as “genoeconomics,” which seeks to identify genetic factors affecting decision-making, and neuroeconomics, which correlates brain activity with decision-making.

Within these two silos, we focus on the emerging literature using experiments with children to understand the determinants of child decision-making, including how economic preferences, information sets, and decision skills affect choices.² Experimental economics has already made a great deal of headway into understanding the economic preferences of adults and linking them to behaviors and outcomes, ranging from health outcomes, such as obesity, to financial outcomes, such as credit card debt (e.g., see Falk et al., 2018). Exploring economic preferences, choices, and relevant trade-offs with children is a natural extension. This exploration aligns with the influential work of Heckman and co-authors in labor economics, who model child skill development with a particular focus on the importance of non-cognitive or “soft” skills (e.g., Cunha and Heckman, 2007).³

With respect to the first silo, children are active participants in their human capital production process, and child preferences, beliefs, and information sets are also inputs into household production and consumption. Measurement of such parameters – and mapping them to behaviors - is therefore key to understanding these foundational processes. Further, intervention studies that seek to improve child education, health, or well-being are numerous. Incorporating

² There are several ways to use children in experiments (as Harrison and List (2004) note, such studies fall under the rubric of their definition of “field experiments”). The first is to use children as subjects to evaluate the impact of various interventions on outcomes (i.e., a random controlled trial, or RCT). Since this is a well-traveled area in economics, we only briefly discuss such work so we can focus attention more immediately on the second way, which is to use experiments with children to understand the development of economic preferences and decision-making (this would typically take the form of studying children in the laboratory or in artefactual field experiments).

³ The work in labor economics tends to focus on skills as measured by various assessments/tests in education, or with respect to “soft skills,” measurement of character or personality using surveys. Different from that work, tools from experimental economics allow economists to measure economic preferences such as e.g., time, risk, or social preferences; or to understand how children react in decision problems.

preference and decision-making experiments into such studies adds value by identifying the role that child preferences play in immediate child behavior (i.e., which interventions will work) and by providing a potentially important variable for heterogeneity, or moderation, analysis (i.e., for whom will interventions work). With respect to the second silo, experiments with children subjects can lend key insights into questions about adults. For example, we can measure the variance in preferences associated with differences in family background at an early age, learn the role of experience in shaping preferences; and estimate when traits and preferences emerge and how they evolve through time.

Why is the study of children by experimental economists necessary, given the substantial body of literature on child development from the sister sciences? First, the parameters that economists are interested in may be distinct from those studied by developmental psychologists. For example, in well-known work starting in the 1960s, psychologist Walter Mischel showed that 3-5 year-old children who can resist eating a marshmallow in order to receive a larger reward have better educational and labor market outcomes than those who cannot (Shoda et al., 1990; Schlam et al., 2013). In our own recent work with a group of 3-5 year-olds, we replicate Mischel's marshmallow paradigm and conduct time preference experiments that are more standard in the economics literature. We find that these two paradigms measure different constructs, and that time preferences are predictive of disciplinary referrals in the classroom years later (Castillo et al., 2020).

Second, economists' training in issues of identification, estimation, and statistical analysis can add value. For example, As Cotton et al. (2020) note, the work of Carrol (1962) and other 20th century contemporary psychologists of education express learning as a function of the ratio of the time spent on learning to the time needed to learn. Cotton et al. (2020) use a field experiment to capture two crucial student-level unobservables—academic efficiency (turning inputs into outputs) and time preference (motivation)—which lend empirical insights into the underpinnings of adolescent skill formation and provide lessons into curbing racial and gender achievement gaps. Likewise, a series of papers have recently emerged suggesting that omitted variables may explain Mischel's results (Kidd et al., 2013). In this article, we discuss how methods from economics can address such identification issues. Finally, experimental economists have developed a toolkit of incentivized, context-free, elicitation tasks that are portable to the sister sciences that explore

complementary behavioral elements, permitting an exploration of economic mediators and moderators.⁴

The remainder of our study begins with a focus on the broader economics literature that uses children as experimental subjects. These studies span a large academic space, from examining children's rationality, to measuring various economic preferences, to an exploration of the variables correlated with children's preferences and behaviors. To place this work in the proper context and aid in its interpretation, we provide a simple conceptual framework. This approach has a dual purpose in that it also permits a deeper understanding of the development of children's economic skills and preferences, and the role of those preferences in shaping behavior.

This exploration in turn informs several fields of economics. For example, experimental economists and economic theorists might find data from children of import to inform preference parameters in economic theories. Labor and education economists might find children data useful to understand the role of children's economic preferences and information sets on behavior that affects their human capital development. Likewise, health economists need information about children's preferences and information sets to inform the design of interventions as well as to study their potentially heterogeneous effects. Finally, researchers studying economics of the household may find value in understanding children themselves as active participants in household bargaining.

We then pivot to a discussion of key identification and methodological challenges the experimental researcher faces when using children as subjects. Our goal is to provide practical advice for conducting experiments with children. For example, how can the researcher most efficiently align child competencies with features of the experimental design? In addition, we discuss moving beyond the traditional A/B experimental approach to lend insights into why certain economic choices are taken. This type of analysis permits a deeper understanding of the behavioral mediators, a necessary step to grasp mechanisms, test theories, and provide efficient policy tools. We next integrate insights from economics and our sister sciences – early education, developmental psychology, and neuroscience – to provide a playbook for how to design and execute field experiments with children. To complement the existing literature, which tends to

⁴ Indeed, a number of recent papers from psychology have moved from using hypothetical decision tasks to using dictator games more standard in the experimental economics literature. These papers are summarized later in this article.

focus on correlational data, we highlight how experiments can be used to aid in identification and interpretation of data. In this spirit, we also discuss practical issues of consent and ethics, setting up experiments, and selection of incentives. Finally, we conclude by offering ideas on how to scale experiments and best practices for documenting experimental procedures with children. We package this section within a catalog of 10 tips for conducting experiments with children. We conclude with thoughts about where the field is headed and suggestions for future directions that we view as most promising.

2. The Growing Economics Literature on Experiments with Children

Experiments with children can provide economists with unique knowledge that cannot be obtained from conducting experiments with adults. This includes studying the development of economic preferences among children to understand adults, as well as using economic experiments to understand children themselves to inform models of human capital formation and early childhood education. Economic experiments allow us to measure aspects of childrens' decision-making such as economic preferences (e.g., risk, time, and social preferences), rationality in decision-making, and information or beliefs that a child possesses. As displayed in Figure 1, the past 20 years has seen an enormous increase in working papers and publications that use economic experiments with children. Given this interest, recent discussions of "alternative subject pools" in experimental economics include children (Fréchette, 2016).

[FIGURE 1 ABOUT HERE]

Some of the earliest work examining children's preferences and decision-making in the economics literature comes from William Harbaugh, Kate Krause, Lise Vesterlund and co-authors. One of the first papers – published in the *American Economic Review* - focuses on investigating whether children are rational decision-makers by examining consistency with the General Axiom of Revealed Preference, or GARP (Harbaugh et al., 2001). The authors motivate their work by pointing out that without an understanding of whether children are rational, we cannot know whether standard economic models apply to them. Harbaugh et al. (2001) find evidence for rationality among children as young as age 7. Harbaugh et al. (2001; 2002; 2003) expand the exploration to other economic preferences across age groups, including altruism, risk preferences, trust, and behavioral biases. The theme of these papers is to compare decision-making by children to the well-documented decision-making patterns of university students in the laboratory.

The literature grew considerably in the mid-2000s, and ever since researchers have simplified standard economic games from the laboratory to facilitate understanding among children, including dictator games, ultimatum games, time and risk preference elicitations, and others. How children behave in these games forms the basis of our understanding of the origins of preferences and decision-making, as well as provides insights for academia and practice.

The Chicago Heights Early Childhood Center (CHECC) is a recent example of a concentrated effort to expand the literature on experiments with children. CHECC is a large-scale field experiment launched in Chicago Heights, Illinois in 2010. We highlight CHECC because the explicit goals closely follow our two bins on the benefits of conducting experiments with children. The first goal of CHECC is to learn about the optimal investment profile in early childhood to reduce academic achievement gaps. To this end, CHECC randomized over 2,000 children to a control group or to an early childhood program targeting parents or children (see e.g., Fryer et al., 2015; Fryer et al., 2020). Findings suggest that targeting parents has a larger impact on executive functions (Fryer et al., 2015) while targeting the schooling input has a larger impact on academic skills (Fryer et al., 2020). Further, findings show that CHECC impacts not only test scores, but brain activity as measured by electroencephalography (EEG). Compared to children in the control group, children assigned to the program targeting children had greater brain activity related to executive functioning, and this brain activity was predictive of executive functioning skills for up to three years after the intervention (Ye et al., 2021).

The second goal of CHECC is to understand broader issues beyond the intervention effects described above. This includes economic preference development, i.e., to understand adults through children. One key to this area of study is understanding how economic preferences, cognitive, and non-cognitive skills serve as inputs into human capital formation (Castillo et al., 2020). Hence, CHECC also operates as an experimental laboratory and researchers invite CHECC children and parents to participate in periodic economic experiments. CHECC also aims to increase the number of economists studying children. Therefore, another focus is to include researchers outside of the original principal investigators of CHECC to conduct studies. To date, experiments that involve understanding the development of economic preferences with children at CHECC have involved nearly 30 different authors and 11 publications. For these reasons, throughout the article we return to studies involving CHECC to augment the learnings from the literature where appropriate.

3. Children can be Rational

Before commencing a discussion of the research agenda into the preferences, choices, and behaviors of children, it is important to understand the make-up of this experimental species. For example, it is interesting, but much less so, to study a species that makes decisions haphazardly, or whose preferences are labile and subject to the whims of circumstance. In such cases, no optimization principles may underlie even the most straightforward of choices (see List, 2002a).

3.1 Rationality and strategic thinking

Rationality and strategic thinking form the basis of most economic theories; hence, understanding their development at an early age has attracted the attention of economists. Expanding on earlier work by Harbaugh and co-authors, recent work has continued studying other violations of rationality (Brocas and Carillo, 2018; Castillo et al., 2018; Brocas et al., 2019), backward induction (Brosig-Koch et al., 2015) and strategic thinking (Sher et al., 2014; Czermak et al., 2016; Brocas et al., 2017; Brocas et al., 2019b). An interesting result is that consistency in decision-making develops asymmetrically across domains. Children first develop consistency in domains that are more familiar to them, for example choosing between toys (Brocas et al., 2016). Because consistency may not be assured, when modeling children's decisions, it is important to test for consistency and control for decision error, especially in examining the effects of economic preferences on field behavior (Castillo et al., 2018).

Concerning strategic thinking, Brocas and Carillo (2018) find that children as young as age 3-5 can think strategically, though they have limited attention, which makes it difficult for them to consistently act strategically. In Sher et al. (2014), researchers find that by the age of 7, children's ability to act strategically by intentionally misleading is similar to behavior of adults. Strategic thinking is related to Theory of Mind (ToM), which is a widely studied construct in developmental psychology that relates to the ability of a child to understand that other people may have different preferences or information states than oneself. The developmental psychology literature suggests that basic ToM develops between the ages of 3-5, but that higher-order ToM abilities continue to develop through early adulthood (e.g., see Dumontheil et al., 2010). A recent paper from economics shows that childhood ToM predicts behavior in economic games and is associated with later-life educational and labor market outcomes (Fe et al., 2020). As such, it is important for economists to understand ToM, and related child traits, not only because ToM is predictive of later

life outcomes, but since models of strategic behavior need to make assumptions about agents' information sets.

3.2 Rationality in Markets

Early field experimental work studying markets includes many child decision-makers in those markets, providing tests of economic theory amongst this group of agents (List and Shogren, 1998; List, 2002a,b, 2003, 2004; List and Millimet, 2008). For example, List (2004) explores whether predictions from competitive models of equilibrium in bilateral negotiating markets accurately predict behavior. Since standard economic theory relies on two assumptions, utility-maximizing behavior and the Walrasian tâtonnement process, relaxing those assumptions remains invaluable. In his lab experiments with Harvard undergraduates, Chamberlin (1948) observes that quantity traded is typically higher and realized prices typically lower than predicted by competitive models of equilibrium. Vernon Smith, an experimental subject in Chamberlin's markets, later refined Chamberlin's work by varying two key aspects of the experimental design in Smith (1962): *i*) use of a double-oral auction where an auctioneer took bids and asks and *ii*) use of multiple market periods, permitting agents a chance to learn rules, strategies, and perhaps even valuations in the market (whereas Chamberlin's markets were one-shot). Empirical results from Smith's experiments are astonishing—both quantity and price levels are close to competitive predictions. This result is an early example highlighting that Walrasian tâtonnement is not a necessary condition in the standard model. By now, hundreds of studies have replicated Smith's original insight.

List (2004) returns to the Chamberlin institution but moves the analysis from the laboratory environment to a natural setting where the actors had experiences with the institution. Much like Smith's (1962) set-up, the market mechanics in these bilateral bargaining markets are not Walrasian. Unlike Smith (1962), however, in these markets subjects set prices as they please, with no guidance from oral bids and asks (a centralized auctioneer). Thus, List's (2004) design shifts the task of adaptation from the auctioneer to the market agents, who mill around making bids and asks in a multi-lateral manner. In doing so, the market structure reformulates the problem of stability of equilibria as a question about the behavior of actual people as opposed to a question about an abstract market. A key result of List's (2004) study is the tendency for exchange prices

to approach the competitive model predictions, especially in symmetric markets (less so in asymmetric markets).

For our purposes, the key data generated in List (2004) are his experiments with children. That is, the experimental markets in which the only participants are children under the age of 13. These include experienced children – who made more than 25 trades per month in the past two years – and inexperienced children, who made less than 3 trades per month in the past two years. The findings are summarized in Figure 2.

[FIGURE 2 HERE]

The patterns are stark, and provide two insights: first, competitive price theory adequately organizes the data when participants have sufficient market experience.⁵ Second, data from the market sessions with purely inexperienced children are akin to a martingale process, whereby prices follow no apparent trajectory or guidance from theory. Inference from these experiments suggest that even in markets populated entirely by experienced children, there is a tendency toward neoclassical expectations after a few periods. Yet, Chamberlin's (1948) original one-shot insights hold in that neoclassical theory does not do well in organizing behavior in markets populated by agents with little market experience. This exploration showcases the value of experiments with children in that they provide an avenue to study behavioral differences across inexperienced agents and their experienced counterparts in a natural manner. We suspect it would be difficult to find a test group outside of children - or visiting villages without markets - that provides such natural variation. More broadly, it is important to understand children's behavior in such exercises because popular explanations of economic behavior model familial decisions as a part of a household bargaining process, in which children are engaged.

The above results are promising since they suggest that with a level of experience children can make rational decisions and think strategically at relatively young ages (or, as in List and Millimet's (2008) tests of GARP, children make a similar level of errors as adults). Researchers wishing to model decision-making by children may consider using standard economic frameworks, however, they should keep in mind other necessary skills that develop concurrently to carry out preferred decisions, such as the ability to reason about others (ToM), attention and impulse control. Hence, before the age of 5-7, researchers who model decision-making by children

⁵ A similar result is observed in Brocas and Carillo (2019a), who replicate List (2004a), finding that 75-82% of groups of 5-11 year-old children converge to equilibrium.

should be concerned about whether it is appropriate to assume standard assumptions of rationality such as transitivity and so on. We discuss the importance, and timing, of child competencies and their implications for design of economic experiments with children in Section 9.

4. Conceptual Framework

Having provided evidence of rationality among children, we turn to a conceptual framework that sheds light on how to interpret data on experiments with children. In sub-section 4.1, we provide a human capital development model to lend insights into the origins and malleability of skills and economic preferences. In sub-section 4.2, we discuss how child skills and economic preferences enter a utility function that influences behavior and decision-making. In our framework, we follow the typical approach, allowing child skills and economic preferences to be affected by investments of parents, schools, and children. Child outcomes are therefore affected by a combination of skills and economic choices, and the latter are partly affected by the child's preferences.

4.1 Development of Economic Preferences

We begin with the assumption that skills and economic preferences result from a nature-nurture (i.e., gene-environment) interaction. In our framework, children are born with a “band” of potential skills or preferences, and the level expressed within that band depends on the environment, including parenting or interventions.⁶ This notion follows the spirit of the genetics literature, which has over the past 25 years moved from a deterministic genetics view (that genetics entirely determine outcomes) toward embracing epigenetics, which posits that gene expression has an environmental influence (Sternberg, 2011; Ijzendoorn et al., 2011; Lickliter and Witherington, 2017).⁷ Related work in the nascent field of genoeconomics seeks to understand the extent of

⁶ Note that our “band” approach is not the only way to conceptualize the idea that genetics and environment both contribute to outcomes (and that there are inherent limits to skill development). We could alternatively model diminishing returns to inputs from the environment on initial skill levels to capture the idea that the environment can affect skills and preferences in a bounded manner.

⁷ The idea of a range of potential outcomes available to each individual – even among animals raised in the laboratory – is due to early work by pioneering geneticist Jerry Hirsch (1963).

heritability of traits. This work suggests that genes explain some - but not all - of the variation in economic preferences and outcomes (Benjamin et al., 2011).⁸

Much of the work that we review in this paper seeks to evaluate the evolution of skills and preferences and to understand what factors affect skill and preference formation. The implicit assumption in this literature is that skills and preferences depend partly on nature and partly on nurture. This means that experiments that seek to understand the inter-generational transmission of economic preferences (discussed in sub-section 6.1) capture both a genetic component and an association with household environment. Experiments that seek to understand the impact of shocks or interventions on outcomes (discussed in section 8) are implicitly measuring the impact of environment on gene expression.

We present a stylized model in which a child's realized skill or economic preference is a reflection of both pre-determined potential skills and preferences and environmental inputs. In the model, we let M_{it} be a vector of skills and economic preferences for individual i at time t . The components of M_{it} could include skills such as cognitive skills, "soft skills," and economic preference parameters representing, e.g., altruism, time discounting, or risk preferences. Let $M_{it} \in [\mathbf{a}_i, \mathbf{b}_i]$, where \mathbf{a}_i and \mathbf{b}_i are vectors representing scalars that indicate lower and upper bounds of skill M for individual i . For simplicity, we assume \mathbf{a}_i and \mathbf{b}_i are known but discuss the relaxation of this assumption later.

In this sense, each component of M is bounded by a separate \mathbf{a}_i and \mathbf{b}_i . Each individual is born with initial conditions M_{i0} where $M_{i0} \in [\mathbf{a}_i, \mathbf{b}_i]$. An individual's lower and upper bound for a particular skill is fixed across time, but the expression of that skill can vary within the bounds at any point in time. Initial conditions M_{i0} are presumed to be associated with genetic background. Papers we review in sections 6.1 and 6.2 seek to proxy for genetic background by studying the association of skills and preferences in very young children with mother's skills, economic preferences, and household demographic/socio-economic (SES) characteristics, respectively.

A distinction between skills and preferences is that for the former, a higher level of skills is typically assumed to be better. Hence, policymakers designing interventions to improve skills may concentrate primarily on \mathbf{b}_i , whereas those seeking to understand the impacts of adverse events on child skills may focus on \mathbf{a}_i . However, for preferences, a normative statement regarding

⁸ Some recent examples are studying the heritability of risk preferences (Linnér et al., 2019) and educational attainment (Lee et al., 2018).

the optimal level of M is in many cases difficult to defend. For example, suppose a_i captures the lower bound on risk preferences (i.e., most risk averse) while b_i captures the upper bound of risk preferences (i.e., most risk seeking). The optimal level is less clear, but may be a function of norms, customs, or local rituals.

The norm or aspirational level of each component of M is captured by M'_{it} . Cultural norms, peers, and parental preferences are a critical aspect of the human development problem. We introduce M'_{it} to capture the idea that each actor may desire for the child's skills or preferences to conform to a group-level norm or a set standard and may incur a cost for deviating from the norm. Specifically, parents choose their level of investment to meet these goals for skills and preferences (Bisin and Verdier, 2001). Policymakers implement their preferred minimum level of skills, by setting educational standards for each grade level and holding schools accountable if they are not met. Children make decisions that are in part influenced by the decisions of their peers (the papers we survey in sub-section 6.3 attest to this). Related work provides evidence for peer and neighborhood effects on child and adolescent behavior (Brock and Durlauf, 2001; 2002; Gaviria and Raphael, 2001; List et al., 2020; see Durlauf, 2014, for a review).

Skill and economic preference development are affected by inputs of the child, parent, and school, prior skill or economic preferences, norms or aspirational levels, individual characteristics, and bounds on skills. Thus, the state variables for the child's development process in period t are:

$$\Omega_{it} = (C_i^t, P_i^t, S_i^t, M_i^{t-1}, M_i'^{t-1}, Z_i, a_i, b_i) \quad (1)$$

where C_i^t , P_i^t and S_i^t capture the history of child, parent, and school inputs (e.g., time spent learning, time with the child, types of activities with which to engage the child) from the beginning of the process up to period t . M_i^{t-1} represents the level of the child's skill and preferences in the prior period, and $M_i'^{t-1}$ represents the level of the norm or aspirational level in the prior period. Z_i are individual characteristics.

The production function of skill and preference development is:

$$M_{it} = f(C_{it}, P_{it}, S_{it} | \Omega_{it}, M'_t) \quad (2)$$

where skills in period t are influenced by inputs from the child, parent, and schools in that period, given the current state and norms or aspirational level. The production function, $f(\cdot)$, is increasing in each argument and quasi-concave. The production of skills and preferences is influenced by skills and inputs from previous periods, and this can reflect self-productivity and dynamic

complementarity, as described in Cunha and Heckman (2007). Self-productivity means that a skill learned in period $t - 1$ persists and can augment learning of that same skill in period t , and dynamic complementarity means that previously learned skills can bolster investment in new skills in period t . The production function reflects the malleability of realized skills and preferences through aspirations and inputs of the child, parent, and school.

Each actor (parents, schools, and children) maximizes their investment in child skills, taking the other investments as given, subject to the production function of skills and preferences (2) and a budget constraint. The budget constraint, $B_{it} = g(p_x, m)$, is a function of prices p_x (where x indexes each activity, such as time spent on the activity or purchases of books or tutoring) and income, m . The outcome of this process yields a policy function of investments that directly relates to the state variables:

$$I_{i,t}^* = G_t(\Omega_{i,t}) \quad (3)$$

We next turn to the implications of this simple framework. The fact that individual bands exist and that they can overlap across people or groups can explain group-level differences that are observed in previous experiments. For example, the literature that we survey in sub-section 5.3 evaluates gender differences in competitiveness between men and women (see, e.g., Gneezy et al., 2003). In much of this literature, women display less competitive preferences than men, which is thought to explain in part the corresponding gap in labor market earnings (Gneezy et al., 2003; Niederle and Vesterlund, 2011). One might conclude from this evidence that the “band” for competitiveness is situated to the right for men compared to women. However, a handful of studies explore behaviors in matrilineal societies and find that, in this context, the gender gap in competitiveness disappears (see, e.g., Gneezy et al., 2009; Andersen et al., 2013). Taken together, these findings are consistent with the idea that the bands for competitiveness overlap for men and women, and we could observe men and women expressing both different and similar M for competitiveness, depending on inputs and environment.⁹

⁹ There is experimental evidence that given the “right” incentives (i.e. the competitive payment is much higher than the sure payment), women compete as much as men (Petrie and Segal, 2017; Ifcher and Zarghamee, 2016). Relatedly, in an exploration of gender differences in social preferences, DellaVigna et al. (2013) find that women are more likely to be on the margin of giving, partly because of a less dispersed distribution of altruism. In the same manner, the distributional aspects of M across groups of interest might hold import for the interpretation of gender competition studies. Similar reasoning holds when interpreting empirical results from interventions when setting M' , which is discussed next.

Another implication relates to the impact of interventions on child skills. Heterogeneity in M_{i0} implies that we may also observe heterogeneity in treatment effects. For example, as implied by Cunha and Heckman (2007), higher M_{i0} could make investments in M more productive. Suggestive evidence for this is found in the intervention discussed in sub-section 2, where the effects of a parenting program were most concentrated among students with high executive function skills at baseline (Fryer et al., 2015). Since M_i is bounded for every child i by a_i and b_i , the implication is that some children may not be able to reach an aspirational level, M' , if it is set above b_i . For such children, interventions intended to increase achievement in a particular skill to reach a goal may not observe considerable effects.

We also find evidence for this result in the CHECC intervention discussed in section 2 – there were no cognitive impacts of the schooling intervention for children who already had high cognitive test scores (Fryer et al., 2020). However, the same intervention could have such an effect on other children if M' is within their band. Aligned with this, the schooling intervention was highly effective for students with cognitive skills below the median (Fryer et al., 2020). A complementary CHECC study that explores the effects of monthly financial incentives amongst adolescents for meeting an achievement standard based on multiple measures of performance reveals overall modest effects but has a large and significant impact among students on the threshold of meeting the achievement standard (Levitt et al., 2016a). Such results imply that intervention success effects will be heterogeneous just by the nature of the policymaker setting M' . In this manner, a natural set of moderators result directly from how the researcher or policymaker sets M' .

Consider an individual's realized skill or preference level, M_i^* . Two individuals with the same level of realized skill could have very different underlying bands of potential ability. In other words, improvement in a skill is different from being born into a “good stock” of ability. This insight poses an interesting question with regards to how early childhood interventions impact later life outcomes. For instance, are long-term outcomes different for children who attained a skill through natural ability or through effort induced by an intervention? This is difficult to tease apart, of course, because a_i and b_i may not be known with certainty. There is some evidence for “reversion to the mean” in the sense that many early childhood interventions see fade-out of skills learned in the years following a program (Duncan and Magnuson, 2013; Almond et al., 2018). This suggests that children who increase their skills through early childhood interventions may

need to be scaffolded with additional interventions in the middle-childhood years to attain persistent gains, whereas those with higher initial skills may maintain their trajectory without additional scaffolding.

The shape of the production function is flexible and, as noted, allows for self-productivity and dynamic complementarities of skills (Cunha and Heckman, 2007). More generally, skills can be complementary, in the sense that improvements or changes in skill X could also lead to increases in skill Y . These complementarities can exist within a period, as well as across periods. In the literature we survey, there is some evidence that skills move in tandem. For example, studies find a correlation between risk preferences and cognitive abilities (Benjamin et al., 2013; Andreoni et al., 2019a). Such correlations also have implications for the identification of skills and preferences, which we discuss in section 10.4.

Finally, the properties of the production function suggest that investing in a child early in life leads to the greatest benefits to the skills within M as skills build on one another, and that investments by more parties (e.g., schools, parents) complement one another. In line with these properties, a review of data from interventions targeted at various age groups have shown benefit-cost ratios that decline with the age of the targeted group (Kautz et al., 2014).¹⁰ Along these lines, List et al. (2018) provide a discussion for the importance in investing in children prior to the school years.¹¹ The implications for economic preferences have not been studied empirically but follow a similar line of thought. Investments early on are more effective for shifting preferences than investments later, and investments in both schooling and parents are complementary. As discussed in section 7, child preferences early on lead to decisions (e.g., conduct at school) and subsequently affect outcomes (e.g., educational attainment). If preferences shift through inputs, then our work at CHECC in which we assess preferences prior to school investments when children are 3-5 is helpful since it may come close to providing evidence of the “innate” nature of preferences.

An open question is to what extent individuals understand the production function or know the values of (a_i, b_i) . The model above implicitly assumes that these are known; however, predictions would change if, for example, individuals had incorrect beliefs about mapping of

¹⁰ However, a newer review did not find this relationship (Rea and Burton, 2020). We view this as an invaluable research question going forward that will be better understood through conducting field experiments with children.

¹¹ Investment in the adolescent years should not be discounted since studies do show impacts of certain investments. For example, a schooling intervention in Joensen et al. (2021) finds that a relatively low-cost ~10 week math class can increase both cognitive and executive functioning skills.

investment effort to outcomes, or about the values of (a_i, b_i) . Examples from empirical work suggest that incorrect beliefs are likely, and it is plausible that academic achievement gaps by SES are associated with differences in beliefs. List et al. (2021) and Boneva and Rauh (2018) find that perceived returns to parental investment are positively associated with household income – i.e., low-SES parents value investment less than high-SES parents. List et al. (2021) describe the existing literature showing that from a very young age, high-SES children do better on academic achievement tests. They conduct field experiments with parents of infants and toddlers in the US to understand whether beliefs can be changed. At baseline, they find that low-income parents and parents with low educational attainment are less likely to believe that their investments in their baby can affect brain development as compared to higher-income and higher-educational attainment parents. The authors also show that interventions consisting of short videos or home visits durably improve beliefs about investments. Similarly, Dizon-Ross (2019) describe a study in Malawi in which giving parents information about their child’s skills affects parental investments.

However, parental interventions do not always affect beliefs. For example, Attanasio et al. (2020) show that mothers of babies in Colombia underestimate the return on their investments, but their home visiting program does not affect beliefs. Further, beliefs are not the only channel through which parental investments are affected: Mayer et al. (2019) conduct an intervention on parents in the US that incorporates reminders, goal setting and social rewards. The authors find that this program affects the time parents spend reading to their children, but that this result does not operate through changes in beliefs.

The concept of mindset in psychology also suggests that people differ in their beliefs about whether they can move M . Children who believe they can change their intelligence (growth mindset) achieve more than children who believe that they cannot do so (fixed mindset) (Dweck, 2015). If actors under-estimate the value of b , or the extent to which M is malleable, this may result in under-investments in child skills. Policies can be designed to counter this issue – for example, Yeager et al. (2019) show that mindset can be changed through interventions.

4.2 How Economic Preferences Predict Behavior

The literature that we survey also focuses on measuring M (section 5) and evaluating how M correlates with contemporaneous or future field behaviors (section 7). Hence, it is helpful to

understand how M enters a child's utility function and leads to changes in behaviors and outcomes. To do this, we provide a standard formulation of a utility function with preference parameters.

Children are assumed to maximize their utility of consumption over time subject to a budget constraint that is a function of prices and income, $B_{it} = g(p, m)$, where income is determined by their parents. Recall that M_{it} can represent skills and preferences. In this case, M_{it} includes preferences over uncertainty, consumption over time, and preferences towards others. The utility function is as follows:

$$U_{it}(c_{it}|M_{it}) \quad (4)$$

where $M_{it} = (r, \delta, \gamma)$, r is a risk aversion parameter, δ is a discount rate, and γ is the degree of prosociality towards others.¹² Utility is increasing and quasi-concave over consumption. Utility has a curvature to account for risk preferences and a discount factor to account for the weight placed on consumption in the present versus the future. While utility is defined over consumption, it could more generally represent field behaviors, e.g., educational achievement, smoking, etc. Choices are affected by these preference parameters in standard ways. More risk averse individuals are less likely to engage in risky behaviors, such as misbehaving in school or smoking. More impatient individuals are more likely to take immediate rewards rather than wait for larger future rewards. Individuals with stronger prosocial inclinations are more generous to others, *ceteris paribus*. To proxy for these underlying preferences, experiments evaluate child and adolescent decisions in allocations between self and others (Section 5.1), choices between safe and risky options and choices between immediate and future rewards (Section 5.2).

Economists who study field behavior often take preferences as given. However, according to our framework, choices are malleable through investments by children, schools, and parents (Section 4.1); this prediction is borne out in the empirical data discussed in Section 8.2. Choices in previous periods may well shape preferences. Children who have fallen behind in school may be impatient and less willing to study to complete coursework that is difficult so that they can graduate. Finally, economic preferences may differ across time because of the development production process and thus yield different outcomes for younger or older children.

¹² Social preferences can include separate components for 'warm glow,' i.e., utility from the gift itself and pure altruism, i.e., utility tied to the consumption of others (Section 5.1). Additional preference parameters that are of interest include present-bias, i.e., the over-weighing of immediate outcomes (Section 5.2), ambiguity aversion, i.e., aversion to unknown risks (Section 5.2) and competitiveness (Section 5.3).

Throughout this section, we have made parallel but separate observations about skills and preferences. Human capital development models focus mostly on skills, whereas models that seek to explain behavior focus more on preferences. An interesting question is to what extent one or the other drives outcomes. One recent paper evaluated skills and time preferences of adolescents, finding that academic efficiency (ability to learn) rather than motivation (time preferences) explains why some students fall behind in school and others do not (Cotton et al., 2020). While the Cotton et al. work represents an early empirical contribution attempting to evaluate the core insights of the contemporary psychology of education models of the 20th century (e.g., Carroll 1962), we trust that future field experiments with children will lend further insights into the underpinnings of adolescent skill formation through the lens of the models of our sister sciences.

5. Can we Measure Children's Preferences?

In this section, we provide a discussion of the types of economic preferences of children that have been explored in the literature, including social preferences, risk and time preferences, and competitiveness. As discussed in Section 4.2, these preferences enter the utility function of the child and influence behaviors and outcomes. A highly recommended companion paper by Sutter et al. (2019) provides a comprehensive survey of experimental findings related to economic preferences with respect to age and gender.

5.1 Social Preferences

All human interaction is founded on our ability to get along with one another. Humans are unique in their ability to form groups and cooperate with genetic strangers. Social preferences – or preferences toward the welfare of others – are a key component of this feature of humanity. Understanding origins of social preferences is important for our understanding of how institutions form and evolve. For example, the large differences in inequality and redistribution between the United States and Scandinavia are mirrored by differences in inequality acceptance (but not in preferences toward efficiency) in cross-cultural experiments among adults (Almås et al., 2019). Indeed, there is a large amount of diversity of social preferences both across and within countries (Falk et al., 2018). Studying how these cross-cultural differences emerge is of interest to the study

of political economy, understanding why people give time and money to charitable causes, and any other economic activity that involves the welfare of others.¹³

To understand social preferences, economists commonly put people in situations where they are given the opportunity to share scarce resources with another person (i.e., a dictator game). At CHECC, in a series of experiments, we gave young children 6 stickers and asked how many they'd like to share with another, anonymous child (Castillo et al., 2020). Figure 3 summarizes these data. Some clear patterns emerge. First, similar to dictator games with adults, we observe a large concentration of sharing at the 50/50 “fair” norm, and children rarely share more than half of their endowment (Panel A). We also observe a concentration in zero sharing—giving none of the endowment to the other person. Yet, there remains marked heterogeneity in allocation decisions. Second, even at young ages, children somewhat increase their sharing behavior as they grow older, from ages 4 through 6.5 years old, both in terms of average amount of endowment shared and the likelihood of sharing at the 50/50 norm (shown in Panel B).

[FIGURE 3 ABOUT HERE]

A similar, and even more pronounced, trend of increasing sharing with age has been documented in a number of studies over the past decade (e.g., Fehr et al., 2008; Angerer et al., 2015a). These studies use dictator games similar to the work at CHECC, as well as allocation decisions with restricted choice sets that allow for direct comparisons of, for example, fairness and efficiency preferences. Studies also have found that trust, reciprocity, and cooperation increases with age (Harbaugh et al., 2003; Sutter and Kocher, 2007; Evans et al., 2013). To study trust, reciprocity and cooperation, researchers typically use simplified variants of commonly used paradigms in laboratory experiments, e.g., ultimatum, prisoner's dilemma, and public goods games (Ledyard et al., 1995). In a field study to explore social preferences of individuals across collector conventions (in public goods games), a television game show (in prisoner dilemma games), and a university fundraising drive (charitable gifts), List (2004b) finds that in every setting there is a strong relationship between age and other regarding preferences—in every case the older subjects act more pro-socially.

¹³ Kenneth Arrow (1972) perhaps put it best when noting: “virtually every commercial transaction has within itself an element of trust... it can be plausibly argued that much of the economic backwardness in the world can be explained by the lack of mutual confidence.”

Rather than showing how behaviors in games change with age, another research approach is to design experiments that allow us to distinguish between theories. For example, List and Samek (2013) study whether dictator game decisions at an early age were due to warm glow or pure altruism, finding evidence for pure altruism but not warm glow. Fehr et al. (2013) explore the evolution of egalitarian, altruistic, and spiteful preferences, finding that altruism became more prominent with age, but that spitefulness and egalitarianism became less frequent with age. Bereby-Meyer and Fiks (2013) find that negative reciprocity increases with age. There are also some instances in which theories apply equally well to children as they do to adults. For example, the theory of moral cost described in Cox et al. (2016), whereby people make decisions based partly on the action set available to them predict the behavior of children as young as 3 as well as the behavior of adult subjects.

Related work also uses social preference experiments with children to understand preferences toward in-group or out-group members, or discrimination. For instance, List et al. (2017) use dictator games that vary whether the recipient appears White or Black and find that 3-5 year-old White and Hispanic children send more to the Black recipient than to the White recipient – a pattern that is reversed in studies of adults (Riach and Rich, 2002). Related work finds that language discrimination in a bilingual city in Italy increases with the child's age, suggesting that perhaps preferences toward minorities develop with age.

An open question revolves around what triggers social preferences to manifest in children. A potential explanation in the case of fairness preferences is that children become more concerned about fairness after they have developed the math skills that allow them to create and comprehend an equal division of the endowment (Chernyak et al., 2016). This underscores the importance of considering cognitive skills that develop concurrently when measuring economic preference development, such within-period complementarities of skills are captured in our conceptual framework of skill formation. Other explanations about why social preferences increase with age include socialization (Fehr et al., 2013) and the ability to take into account the preferences and viewpoints of others (Bereby-Meyer and Fiks, 2013).

5.2 Time and risk preferences

Time and risk preferences play an important role in a wide range of decision-making contexts, which are of interest to economists in fields that span from household finance to health

to education. To measure time and risk preferences, economists usually put people in situations where they make decisions about receiving rewards sooner versus later, and about receiving sure rewards versus being entered into a lottery for rewards, respectively.

Research shows that time preferences, and especially present-biased preferences, are related to participation in financial education programs and credit card borrowing among adults (Meier and Sprenger, 2010; Meier and Sprenger, 2013). These preferences are also related to health outcomes – for example, impatience is correlated with higher BMI (Ikeda et al., 2010; Sutter et al., 2013; Golsteyn et al., 2014; Weller et al., 2008) and more smoking, drinking, and drug abuse behaviors (Bradford et al., 2017; Chabris et al., 2008). Risk tolerance is also linked to smoking and drinking (Anderson and Mellor, 2008). Finally, time preferences are linked to misbehavior, educational attainment, and lifetime income (Castillo et al., 2011; Castillo et al., 2019a; Golsteyn et al., 2014).

Understanding time and risk preference development is important for several reasons. First, we want to understand the origins of such preferences, and how they are shaped by parents and schooling. Second, we want to understand time and risk preferences of children to inform policy interventions. For example, most studies find that patience increases with age (see Sutter et al., 2019, for an excellent overview). This suggests that interventions targeting children should focus more on the present than on the future – i.e., when giving incentives for reaching education or health goals, these incentives should be distributed as soon as possible (Levitt et al., 2016b). Alternatively, adults may be more willing to accept delays in receipt of their incentives. As another example, researchers worry about the apparent greater risk tolerance of adolescence that can lead to detrimental behaviors like smoking and drinking, yet experiments show us that we should be most worried about their tolerance toward ambiguity, i.e., toward unknown risks (Tymula et al., 2012).

Using the CHECC data, Castillo et al. (2020) find that time preferences measured at age 3-5 are associated with disciplinary referrals years later in elementary school. Importantly, these preferences are distinct from cognitive skills (e.g., reading, writing, math) and executive functions (e.g., inhibitory control and working memory) in that they have an independent role in predicting disciplinary referrals. Hence, Castillo et al. (2020) refer to economic preferences as a key previously missing factor in our understanding of human capital formation.

Figure 4 shows the raw data from time and risk preference experiments we conducted at CHECC when children were 3-5 years old. In time preference experiments (Panel A), we gave children choices between smaller, earlier candy rewards and larger, later rewards.¹⁴ In risk preference experiments (Panel B), we gave children opportunities to choose between smaller, certain rewards or different probabilities of larger rewards.

[FIGURE 4 ABOUT HERE: HISTOGRAM OF CHECC DATA]

In Figure 4, we immediately observe heterogeneities in economic preferences, even at these young ages. This lends empirical support for the conceptual framework in Section 4. There is a range of behaviors manifested by the children. How these heterogeneities arise, and how malleable they are, remain open questions for future research.

5.3 The Competitiveness Gender Gap

A concern in the field of labor economics is that in many developed countries women are pursuing higher education at similar, or higher, rates as men, yet women continue to have relatively lower earnings and are less likely to hold executive positions in firms (Bertrand, 2009; Blau et al., 2010; Bertrand and Hallock, 2001; Wolfers, 2006). Gender differences in competitiveness have been proposed as one reason for this gap, and a large literature has used experiments to show that men and women have heterogeneous preferences over wage contracts, with men preferring relative wage incentives more than women (Gneezy et al., 2003; Niederle and Vesterlund, 2007; Gneezy et al., 2009; Flory et al., 2015; Niederle and Vesterlund, 2011, provides a review).

For their part, experiments with children have begun to inform us when this gap develops and what we can do about it. In an early paper on this topic, Gneezy and Rustichini (2004) asked Israeli children to run a short distance either alone or in competition. The authors find that competition improved the running time of boys, but not of girls. The work that followed has mostly preferred a choice-based measure of competitiveness preference rather than a performance-based measure. Most of this work followed Niederle and Vesterlund (2007) by asking children to perform a task under a piece rate and under a tournament, and then eliciting a choice of piece rate or tournament to assess willingness to compete.

¹⁴ Children also completed the “marshmallow” task (Mischel, 1989). In Castillo et al (2020), we show that the time preference task and the marshmallow task are distinct drivers of behavior and only time preferences were predictive of disciplinary referrals.

An overview of the findings in this literature is that competitiveness increases as children get older, and that girls are generally less competitive than boys (for an example, see Sutter and Glätzle-Rutzler, 2015).¹⁵ However, there are some important additional take-aways. First, culture seems to have a large influence. The gender gap in the running task is not observed among Swedish children or Armenian children (Dreber et al., 2011; Khachatryan et al., 2015), in contrast to Israeli children (Gneezy and Rustichini, 2004). Nor is there a gender gap in ball throwing in a matrilineal society in India (Andersen et al., 2013). And, girls from same-sex schools are as competitive as boys (Booth and Nolen, 2012). Second, task differences also seem to be important. For example, the gender gap in competitiveness is not observed in the “female” sports of skipping rope and dancing (Dreber et al., 2014; Khachatryan et al., 2015).

The literature suggests that culture, norms, and social learning play critical roles in the development of competitiveness preferences. Cultural norms of behavior may impose rigidities on the expression of skills so that patterns of gender differences in competitive preferences seemingly reverse (see Andersen et al, 2013, as an example). Culture and norms may also play a role in time, risk, and social preferences, where there is as much diversity across countries as within (Falk et al, 2018). This highlights the import of task selection (and potentially needing many types of tasks) to elicit a competitiveness trait. In light of the model discussed in Section 4, the broad literature is consonant with the notion that the “band” of competitiveness preferences of boys is to the right of that of girls, but there is an overlap, and culture and norms affect whether or not there is a manifested gender gap in competitive preferences. For instance, while most boys might be more competitive than most girls, different influences from culture or parenting can cause girls and boys to look similar (i.e., end up in the overlapping part of the band) or reverse the identity.

6. What Variables Correlate with Children’s Preferences?

A key to understanding the causes of children’s behavior is having a firm grasp of their preferences, not only the origination of such preferences but also the correlates. Children can be influenced by a myriad of factors, from their parents to their neighbors to their environment. This

¹⁵ Yet, the work of Flory et al. (2018) shows that while research on competitiveness has emphasized sex as a physiological determinant, their evidence shows that women's preferences over competition change with age such that the gender gap, while large for young adults, disappears in older populations (because older women are much more competitive). This result suggests that a simple gender-based view of competitiveness is misleading, and that age seems just as important as sex. Future field experiments with children in this space can bear valuable fruit, as early life events will surely be meaningful.

section explores correlates with children’s behaviors from a broad variety of empirical approaches but focuses on studies using field experiments, as defined in Harrison and List (2004).

6.1 Children and their Parents

Children spend a disproportionate amount of time with their parents, especially when they are young. As described in our model in Section 4, the influence that parents have on their children could affect their skills and economic preferences, for their patience, willingness to take risks, competitive attitudes, and prosocial behavior. The literature on improving the outcomes of disadvantaged children has started to focus on parental investments (see List et al., 2018, for an overview). Fryer et al. (2015) uses the CHECC program to explore the impact of intervening with parents on improving child non-cognitive skills.

Researchers hypothesize that parenting shapes economic preferences (Bisin and Verdier, 2001; Lundberg et al., 2009; Doepke and Zilibotti, 2017; Del Boca et al., 2019). This could occur as the child grows and decisions shift from being made by the parent to being made by the child. During this time, parents transmit their preferences to the child, whether their child might agree with their parent or not (possibly through parental investments in the child, see our model in Section 4). Understanding if and how this occurs requires an empirical evaluation of children and their parents.

Certain evidence shows that parents and their adult children share similarities in economic preferences (Dohmen et al., 2011; Kimball et al., 2009). However, research with parents and children finds mixed evidence of a correlation for patience (Bettinger and Slonim, 2007; Kosse and Pfeiffer, 2012; Andreoni et al., 2017; Samek et al., 2019) and no evidence for a correlation in altruism (Ben-Ner et al., 2017; Bettinger and Slonim, 2006). There does seem to be a correlation for risk preferences between children and parents (Alan et al., 2017; Chowdhury et al., 2018; Samek et al., 2019).¹⁶ Further, Khadjavi and Nicklisch (2018), find that parents’ ambitions for their children are correlated with child competitiveness. We interpret the literature as suggestive evidence that parents do indeed transmit their preferences to their children, and the finding of a more persistent correlation among adult children than among young children suggests that many years of parental socialization are required for preferences to “stick”. An alternative explanation

¹⁶ Alan et al. (2017) study mothers and the correlation is only present for daughters.

is that the correlation exists even at young ages but that preferences are difficult to measure – or are noisier - in early childhood.

A related area of study is how children affect decision-making in the household and how that decision-making affects a child's choices. A child's preferences interact with her parent's preferences, and this can affect many aspects of behavior within and outside of a household, such as human capital development of children, products purchased in the household, and constraints parents might place on a child's time. The interaction between parent and child within the household is typically modeled as a non-cooperative game and preferences are inferred from survey data or parameterized from auxiliary experimental data sets (Lundberg et al., 2009; Del Boca et al., 2019).

Examining the interaction of children and parents in decision-making, Houser et al. (2015), using children from CHECC, experimentally examine the effect of having a child present when parents make a decision whether to cheat. They find that in the presence of girls – but not of boys – parents are less likely to cheat when their child can observe their behavior, especially when the gift is for the child. In another CHECC study, Ben-Ner et al. (2017) examine the impact of children on donation decisions of their parents. In this study, parents are asked to make a donation decision that in some cases is observed by their child. The authors find that some sub-sets of parents (fathers, and parents of generous children) are more likely to model generous behavior when they know that their child will see their decision. These results suggest that parents do make decisions aimed at influencing children's preferences as we predict in our model in Section 4 – for example, modeling honest or generous preferences in a potential attempt to encourage their children to adopt such preferences.

6.2 Children and their Socio-Economic Status

Inequality is an important societal issue that is of interest in all fields of economics. What explains inequality, and how to reduce inequality, are open questions. In light of our model discussed in Section 4, intervening in childhood can lead to changes in the outcomes within the “bands” of children from different socio-economic backgrounds. For example, Dynarski et al. (2018) conduct a field experiment to evaluate the impact of information and promise of financial assistance upon admission on encouraging low-income, high-achieving students to enroll at the

University of Michigan. The number of students who apply and enroll doubles, suggesting such an approach could have strong impacts on closing income gaps through college choice.

One possibility is that inequalities are persistent due to differences in economic preferences between different populations. Indeed, experiments with children show associations of SES with economic preferences. Research shows that children and adolescents from low-SES backgrounds tend to be less patient and more risk seeking than higher-income counterparts (Eckel et al., 2012; Deckers et al., 2015; Falk and Kosse 2016). With one exception (Almås et al., 2016), most studies also find that lower SES children are less altruistic/egalitarian (Benenson et al., 2007; Bauer et al., 2014b; Deckers et al., 2015; Falk and Kosse, 2016; Kosse et al., 2019).

The findings are relevant since exhibiting more impatience as a young child or adolescent could have compounding effects on educational choices - e.g., misbehavior in school or graduating from high school (Castillo et al., 2011; Castillo et al., 2019a; 2020). Adolescents who are more tolerant to risk may take up drugs or other risky behaviors, which could also impact their future outcomes (Rao et al., 2011). It is less obvious how other preferences – e.g., social preferences – affect inequality. However, as children become adults and start making decisions that affect institutions, more egalitarian preferences in society can lead to different policy decisions (e.g., with regards to taxation and redistribution) (Cappelen et al., 2020).

An important caveat is that we should not draw causal inference from the studies described above. It is not clear whether differences in SES cause the differences in preferences, or whether omitted variables drive these differences. Omitted variables could include differences in access to resources, differences in parenting, or differences in cognitive abilities. Incorporating these variables into the analysis will help to limit (but not fully solve) this problem. Another open question is whether differences in preferences by SES are present at birth, or whether interactions with the environment result in these differences. There is some evidence for the latter in a study by Rao (2019) -- the economic preferences of adolescents from higher income backgrounds are affected by random assignment to classrooms with adolescents from lower income backgrounds.

6.3 Children and their Peers

In the framework developed in Section 4, we discuss a group-level norm of M' that incentivizes the child to conform to the norm. To this end, related experimental literature evaluates the impact of peers and neighborhoods. In List et al. (2020), the authors take advantage of the

CHECC field experiment to explore the social side of human capital formation and skill development. While most of behavioral economics is founded on psychological concepts, List et al. (2020) leverage insights from sociology (for example, Coleman, 1988) to examine neighbor and peer effects on human capital formation of 3-5 year-olds. They can make causal statements on neighbor effects because CHECC randomized treatment allocation across space and time, effectively creating a checkerboard panel data set to exploit. List et al. (2020) document large localized spillover effects on control children who live near treated children. For our purposes, perhaps the most compelling evidence that speaks to our theoretical framework is the underlying economic mediator (mechanisms) at work. They find that the spillover effect on executive function scores operates through the child's social network while parental investment is an important channel through which cognitive spillover effects operate. Both results highlight that human capital accumulation is fundamentally a social activity, even at very young ages.

Peer effects are also explored in Rao (2019), who examine the effects of a change in school policy in India whereby children from poor families are placed in classrooms with rich families. Rao (2019) collects administrative and experimental data on affected children and finds that children from rich families are more prosocial after having been exposed to children from poor families. By contrast, there are limited impacts of such exposure on academic achievement.

Another approach to understand the impact of peers is to conduct manipulations in which child decisions are either made public or private. This is most commonly done to study social preferences. In work that uses this approach, children as young as age 5 are more generous when the recipient is aware of their donation options (Leimgruber et al., 2012). Similar results are found in Houser et al. (2012) for children aged 9 and above when their decisions are made in public environments, and in Chen et al. (2016) for children aged 8 when their choices are displayed to their peers. Children also tend to be more generous toward in-group than out-group members, see for example Bindra et al. (2018) for a study with children ages 3-6, Bauer et al. (2014a) for ages 7-18, Sparks et al. (2017) for ages 4-6, and Zhu et al. (2015) for ages 2.5-6.5.

Related work finds some evidence for peer effects on other economic preferences of children, but this work is limited. Eckel et al. (2012) find that a higher percentage of low-income peers in a school is correlated with greater risk aversion among adolescents. However, this paper does not seek to identify whether this was a causal effect. Charness et al. (2019) find that prompting children to think about peers improves their performance on a ToM task.

7. Do Preferences Correlate with Field Behavior?

A large literature evaluates the association of adults' economic preferences with field behaviors. This work finds that economic preferences are associated with field behaviors in ways that are consistent with the framework in Section 4.2. For example, increased patience is associated with savings behavior, lower credit card debt, and improved health outcomes, which are also associated with what we view as positive life outcomes (Chapman and Coups, 1999; Picone et al., 2004; Weller et al., 2008; Bradford, 2010; Meier and Sprenger, 2010; Leonard et al., 2013). Research with children and adolescents finds similar associations. These results are important since they suggest that economists are measuring a construct that also has field relevance among children. Importantly, however, observing preferences and behavior at the same time does not speak to causal effects as preferences and behavior may co-evolve.

The literature does teach us that time and risk preferences of children and adolescents are correlated with smoking, drinking, body mass index (BMI), savings, and conduct at school some of the time (e.g., Sutter et al., 2013). Focusing on a small sample of truant adolescents, risk and time preferences are not found to be correlated with school attendance (Antrobus et al., 2017). Samek et al. (2019) find an association of time preferences with BMI of adolescents when time preferences are measured using a stated-preference survey, but not when preferences are measured using an incentivized experimental elicitation task. In another paper, competitive preferences are correlated with adolescent choices of academic track in school (Buser et al., 2014).

Increasing the distance in time between the measurement of preferences and field behavior allows for an examination of the significance of early preferences on future outcomes. Issues of causality remain relevant – for example, other co-evolving omitted variables could drive the associations, increasing causal density. This problem can be mitigated somewhat by including a large number of covariates as control variables. Finding that measured preferences earlier in life are correlated with longer-run behavior confirms the importance of understanding children to understand adults, yet in the end the results remain correlational.

In a nutshell, however, the literature does show that time and risk preferences measured in adolescence are correlated with educational and labor market outcomes later in life. And, more impatient and risk-taking adolescents, as measured with incentivized experiments, are more likely to misbehave in school (Castillo et al., 2011; Castillo et al., 2018) and are less likely to graduate from high school five years later (Castillo et al., 2019a), even controlling for demographics,

cognitive ability and school effects. Other studies using unincentivized measures find a relationship between present-biasedness, self-control and impulsivity in children, adolescents and young adults and labor market and health outcomes (DellaVigna and Paserman, 2005; Moffitt et al, 2011). Golsteyn et al. (2014) use a hypothetical time preference measure with teenagers and find a relationship with labor market outcomes 20 years later. Cadena and Keys (2015) use a measure of impatience from an interviewer report of whether the young adult is “impatient or restless” during the interview and find that it is related to school drop-out and labor market outcomes later in life.

Examining behaviors at even younger ages, we find that there is evidence for preferences playing a role in life outcomes as well. Mischel et al. (1989) show that a child’s ability to resist eating a marshmallow at age four predicts social and cognitive competency during adolescence (see also Mischel et al., 1988; Shoda et al., 1990). However, some scholars challenge the robustness of these findings with additional data (Watts et al, 2018; Falk et al, 2019). In particular, the marshmallow paradigm may not be a clean measure of the ability to delay, since manipulating the uncertainty in the environment also has large impacts on delay decisions (Kidd et al., 2013). Further, Benjamin et al. (2019) collect data from the original Mischel and Ebbsen (1970) sample 40 years later and find no effect of the duration of time waited on a variety of life outcomes, but they do find an impact of an index of self-regulation measured at various later ages on adult outcomes. This study suggests that the marshmallow paradigm could also be picking up aspects of the child’s environment or is less informative for later life outcomes than advertised in the literature and popular press. It also points again to the importance of considering omitted variables as discussed in Section 6.2.

Exploring behaviors in CHECC children, we find that child time preferences as measured by incentivized experiments at age 3-5 explain disciplinary referrals in elementary school up to six years later, even when controlling for cognitive abilities and executive functions (Castillo et al., 2020). Importantly, we find that economic preferences are a third, distinct factor shaping life outcomes (in addition to cognitive skills and executive functions).

8. Identification of Preference Development

Discovering causal relationships is invaluable since they represent the key to theory testing, the basis of fact, and are vital for informed policy prescriptions. For example, one important

exercise is to identify the impact of the inputs – child, parent, and school – on human capital formation, as discussed in Section 4. Using naturally-occurring data (in many cases “natural experiments”) and invoking various assumptions represents the standard approach to provide such causal estimates in economics. However, the economics community has turned increasingly to using random assignment, or field experiments with well-specified models.¹⁷ For some policy questions of interest, this approach can be challenging to implement or potentially unethical or ill-advised. For example, to test the effects of domestic violence or low SES on a child’s economic preferences, we would not want to force individuals into disadvantaged circumstances. Instead, interventions should have the intention of delivering positive change. For example, random assignment to income transfers to increase the household budget would be an appropriate way to experimentally measure the effects of a change in income. Alternatively, an intervention to reduce household violence could be used to evaluate the effect of violence on child skills and preferences.

There is a large and growing literature on the effects of early childhood, or in utero, environmental effects on child cognitive skills and misbehavior (Almond and Currie, 2011). Yet research on the effects of interventions, natural experiments, environmental conditions or negative shocks on economic preferences is still in its infancy. Our model in Section 4 suggests that it is also important to understand these effects since we consider preferences malleable. Interventions aimed at improving household financial wellbeing – such as cash transfers – or interventions aimed at improving child cognitive skills through educational programming – may have unintended (positive or negative) consequences on economic preferences.¹⁸ In turn, preferences shaped in childhood may affect individual outcomes in adulthood, as well as voting and decision-making that potentially impacts institutions more broadly.

When measuring contemporaneous environmental factors and economic preferences without relying on experimentation, there can be a problem of reverse causality. A good example of a reverse causality problem is in the nutrition field, which reports that restrictive parent feeding practices are positively correlated with childhood obesity (Shloim et al., 2015; Boswell et al., 2019). Is it the case that obesity is caused by restrictive feeding practices, or that obesity causes parents to restrict feeding? Outside of an experimental evaluation, one solution is separating the

¹⁷ Fryer (2017) provides a review of randomized field experiments in human capital formation.

¹⁸ There is a large literature on the effects of conditional cash transfer or support programs on a variety of life outcomes (e.g., Hoynes et al., 2016 on SNAP in the U.S.; Attanasio et al., 2012 and Parker and Todd, 2017 on PROGRESA in Mexico)

evaluation in time as is done in several papers in Section 7 – e.g., measuring feeding practices at $t=0$ and returning a year later to measure BMI at $t=1$. This separation in time must still address the omitted variables problem discussed in Section 6.2, since other factors could be driving the correlational effects observed. While one could strive to incorporate all manner of observable characteristics as controls, other unobserved drivers of behavior are still possible. This points again to the benefits of randomization for identification.

8.1 Exploiting natural experiments to understand preferences

Economists have been exploiting natural experiments for decades. A common example of a natural experiment uses changes in compulsory schooling rules to understand the returns to education (e.g., Pischke and von Wachter, 2008, Bhuller et al., 2017). Natural disasters, wars, bombings, and civil conflict are other exogenous events that can have a profound effect on outcomes and preferences. Bauer et al. (2014a) show that children exposed to war become more egalitarian towards their in-group. Castillo et al (2019b) find that being exposed to political violence while in utero produced a higher willingness to take risks 20-30 years later as an adult, and Castillo (2020) finds children exposed to domestic violence are more risk averse at age 12. Callen et al. (2014) examine how psychological priming interacts with exposure to violence 0-8 years earlier on risk preferences. Adults exposed to violence up to 8 years earlier and primed with fear were more likely to prefer certain outcomes relative to uncertain outcomes.

Income shocks can also be used as natural experiments. Carrillo (2020) examine the variation in coffee cultivation in Colombia during children's schooling years and find that cohorts that faced higher returns to coffee-related work in childhood completed less schooling and had lower adult earnings. He suggests that these findings are consistent with the possibility that students ignore or heavily discount the future when faced with immediate income gains.

Changes in school policies and assignment to classrooms and schools may provide interactions with new groups of individuals. We discussed studies that take advantage of these new interactions to study peer effects in Section 6.3 (List et al., 2020; Rao, 2019).

8.2 Interventions and effects on preferences

The literature on interventions to understand the stability and malleability of preferences is of two types. The first type develops a program that specifically targets a particular economic

preference and evaluates the impact of the program on that preference. For example, Alan and Ertac (2018a) find that a curriculum targeting forward-looking behavior and imagining a future self increases patience in incentivized intertemporal choice experiments. Lührmann et al. (2018) find that a financial literacy program increases time consistency. Alan et al. (2019) find that an educational intervention teaching grit increases persistence in a real effort task, improves math test scores and reduces the gender gap in competitiveness (Alan and Ertac, 2018b). Based on the framework in Section 4, such interventions work by changing the inputs of the schools, S .

The second strand of this literature evaluates the impact of interventions with alternative aims – e.g., programs aimed at reducing the academic achievement gap – on economic preferences. For example, the CHECC program was designed to impact educational attainment through early interventions with children and parents. This included a Parent Academy program that targeted parents (through classes on how to teach to children at home) and a Preschool program that targeted children (by providing a free, full-day high-quality preschool). Because these programs affect inputs P and S , respectively, they might also affect economic preferences. For example, the Parent Academy program might affect children through changing how parents interact with their children or how they model behavior, while the Preschool program might affect children through exposing them to interaction with peers and teachers.

Cappelen et al. (2020) report causal evidence of the impact of these programs on social preferences. They find that the CHECC Parent Academy increases the relative weight children place on efficiency versus fairness, while the Preschool makes children more egalitarian in their fairness views, relative to the control group. To explain these results, Capellen et al. (2020) suggests that parents modeled more efficient allocations in the Parent Academy treatment, and that the Preschool exposed children to sharing rules. A related study by Kosse et al. (2019) finds that a mentoring program aimed at enriching the social environment for children ages 7-9 increases prosociality by 27% of a standard deviation relative to the control group. Similarly, Bettinger and Slonim (2006) examine the effect of school vouchers on prosociality of children (ages 8-16 years old). They find that children in families that received an educational voucher are more generous when given the opportunity to make donations to a charity but are no more generous when given the opportunity to share with a peer.

There is little evidence on the long run effects of interventions on preferences. It is important to collect long-run data because such effects may emerge later in life. For example,

displacement effects of moving out of public housing in Chicago due to demolition show minimal short-run effects (Jacob, 2004), but have large positive effects on long-run outcomes, such as employment and earnings (Chyn, 2018). Research shows that the effects of early childhood interventions tend to fade out in the intermediate term, yet the Perry Preschool project has positive impacts on treated children in adulthood, and further, this program has intergenerational effects on the children of treated children (Heckman and Karapakula, 2019).

Studying moderators, or the heterogeneous impact of an intervention, benefits from understanding the children themselves. For example, Fryer et al. (2020) show that the CHECC preschool is most beneficial for children with below-median cognitive and executive functioning skills. A related area is to explore how economic preferences interact with an intervention. One example of such a study is Datar et al. (2019), who find that adolescent time preferences mediate the impact of environment on obesity. There are additional open questions related to interaction effects. For example, what determines parental investments in a child and how might that interact with child behaviors or preferences? Also, how much does the mapping (or returns) of parental investment vary across children?

9. Child Competencies and Experimental Design

The disciplines of developmental psychology, education, and neuroscience provide insights into the child development process, but the design features of economic experiments are unique and require further interpretation. The rapid increase in experiments with children is promising, but it is important to step back to consider how best to design experiments for every age range. This can reduce noise in the data and make comparisons across experiments more useful. While we cannot answer what would constitute the best practice yet, we make steps in this direction in several ways.

First, we discuss the development of child competencies taken from the sister sciences – developmental psychology, education, and neuroscience -- that are important to consider when designing economics experiments. Second, we summarize the methods used in experiments with children of different ages to date, including organizing our thoughts around what we think are the key design and protocol considerations and synthesizing the above to provide suggestions for how competencies can be taken into account when designing experiments. To assist in this task, we must integrate an understanding of when competencies develop (from the developmental

psychology and cognitive neuroscience disciplines) with an understanding of what educational tools are effective for explaining concepts to children at different ages (from the education discipline).

9.1 Child Competencies to Consider in Design

Two core child competencies that we must consider for the design of experiments are academic skills (e.g., literacy, numeracy) and executive functions (working memory, inhibitory control, attention shifting). Additional aspects of child development that merit consideration are abstract thought, the ability to take the viewpoints of others into account, and fine motor skills. Our goal is to align the design of the experiment with these competencies.¹⁹ If a child is unable to process a concept or do a calculation required, this will generate noise in the data and make it difficult to assess the underlying preference; in fact, in such situations no optimization principles may underlie even the most straightforward of choices, rendering an experimental exercise difficult, if not impossible.²⁰

Here, we outline 5 competencies that we consider relevant for conducting economics experiments with children. We focus on the development of competencies from ages 3-18.²¹

1. Numeracy: Numeracy is the ability to understand and work with numbers. Foundational numeracy allows children as young as age 3 to compare groups of items and understand whether groups contain equal or different amounts of numbers. Formal set knowledge – i.e., counting each item only once, counting in order, and stopping at the end of the count – develops later in life (Gelman and Gallistel, 1978). Before young children have set knowledge, they may use shortcuts to divide items, for example by using a turn-taking heuristic to split a pile of prizes equally between two people.

¹⁹ While we focus attention on traditional experimental approaches with children, it is instructive to consider the work of Ye et al. (2021), which highlights the demands when using electroencephalogram (EEG). In this case, the situation demands that the child sits still for a certain amount of time with the EEG nodes on, is not easily distracted by noise such as traffic, fully understands the task, etc. If using children that are too young, the attrition in such work will be substantial, therefore limiting the usefulness of data collected since the discard rate is so high. Broadly, this study highlights, in an extreme case, the importance of aligning child competencies to the demands of the experimental task.

²⁰ An exception to this is if the goal of the study is to examine the ability to perform a complex task.

²¹ Other social scientists have developed methods to evaluate children at even younger ages. For example, eye tracking and measures of attention can be used with infants (e.g., Meristo and Surian, 2013). However, beyond work with parents, and how their inputs into infants map to certain outcomes (such as, e.g., List et al. 2018 and Leung et al, 2020), we did not find any economic experiments using infants, therefore, we limit our discussion to age 3 and up.

The take-away here is that experiments with young children should be designed to appeal to foundational, and not set, knowledge. For example, by age 3, children should be able to identify that larger amounts of rewards are better than smaller amounts, which is of great import because maintaining control within the experiment requires that the reward medium is not only salient, but that utility is increasing in the reward (Smith, 1982). At the same time, they may not be able to accurately count how many rewards they have received. Hence, in a dictator game, rather than asking children to report how many of a set they want to share with another person, experimenters should give children a plate of items and tell them to divide them as they wish, without using numbers. This general approach helps to achieve dominance within the experiment, a condition whereby the reward structure dominates the subjective cost of the activities of the experiment (Wilde, 1980). Further, the literature teaches us that using familiar objects that are similar to one another, and arranging objects in lines with similar spacing, can help children with making comparisons (Carper, 1942; Potter and Levy, 1968; Tan and Bryant, 2000; Clements et al., 2019).

Importantly, one must distinguish the ability to count from set knowledge – just because children can count, it does not mean that they have set knowledge that allows them to count items in an experiment accurately. Set knowledge can be tested by asking children to share “X items” and observing whether the correct number are shared. As children get older, they evolve from being able to count 2-3 items reliably, to being able to count more. Set knowledge (up to 10-20 items) and basic addition and subtraction are part of the core mathematics standards for 5-year old Kindergarteners in the U.S., and therefore by age 6, children should become more savvy in their ability to manipulate up to 20 items in an economics experiment (California Common Core State Standards, 2013). By 1st grade – around age 7 – students should be able to extend this knowledge to 100 items. Note that addition and subtraction lags counting – children can typically count to higher numbers than they can add or subtract (Zur and Gelman, 2004; Campbell, 2005).

The next numeracy milestone relevant to economics experiments is understanding multiplication, division, and fractions. This more formal knowledge is useful for e.g., explaining public goods games that vary the marginal per capita return (MPCR) that determines payoffs. The formal understanding of multiplication and division is expected by 3rd grade – around age 9. But, the next several years are spent on solidifying this knowledge and learning to perform more complex manipulations of fractions. However, even if children do not have a formal understanding of division yet, they are able to understand how a set of physical items can be distributed among

themselves and others. For example, they may use shortcuts like turn-taking heuristics to split a pile of prizes equally between people. Hence, using physical props and familiar language can overcome this hurdle.

The final useful numeracy concept is understanding of probability. In the U.S., the common core math standards expect this concept to be formally introduced in 6th grade – around age 12 – with more in-depth statistics instruction following in grades 7th and 8th - age 13-14 (California Common Core State Standards, 2013). However, children at younger ages than 12 are exposed to – and likely understand – the basic concept of probability. Efraim Fischbein is credited with early work on the development of probabilistic thinking. His novel experiments evaluated the ability of children as young as age 3 to reason about probabilities and fractions (Fischbein, 1975). His work suggests that intuitive reasoning about probability and chance develops early, and he finds that in some cases children ages 3-5 perform better than children ages 12-14 when evaluating the chances of events (Fischbein, 1975). Further, children older than 14, as well as adults, may have difficulties with probability concepts. For example, research documents persistent over-estimation of small probabilities and under-estimation of large probabilities among adults (Kahneman and Tversky, 1979; Burns et al., 2010).

A potential path to harness intuitive understanding of probability in economics experiments is to use spinners – wheels with differently colored slices -- to explain concepts of risky options. This is done in the early work of Harbaugh et al. (Harbaugh et al., 2002) and in many other experiments that we reviewed. Physical examples that are used in laboratory experiments with university students can be applied to experiments with children. For example, Halevy (2007) uses bins with balls to explain ambiguity and risk in laboratory experiments. The use of physical bingo cages is also popular in laboratory experiments and can be used with children (Harbaugh et al., 2002; Harrison et al., 2007; Tanaka et al., 2010).

2. Literacy: Literacy is the ability to communicate using language. For researchers working with young children, an important distinction is productive vocabulary (spoken language) versus receptive vocabulary (understanding when spoken to). Receptive vocabulary develops sooner than productive vocabulary, meaning that at any given time, children can understand more than they can speak. Young children who cannot speak many words may still be able to respond to the experimenter through pointing and other gestures. This represents a key impetus as to why we suggest one-on-one implementation and providing physical props for children below age 5.

Another experimental milestone in children is reading and writing, since this allows researchers to give subjects the independence to read instructions and mark their preferences. Most children learn to read by age 7, with fluent and independent reading emerging around age 9 (Morphett and Washburne, 1931; Fitzgerald and Shanahan, 2000). Unlike with numeracy, there is no intuitive reading or writing ability that emerges earlier than this, though of course some children master reading and writing at earlier ages. We therefore do not advise conducting experiments that require substantial reading or writing with children below the age of 10.

The good news is that there are many alternative options. One alternative is instructions that involve pictorial representations (e.g., pictures of smaller amounts of candies today and larger amounts of candies tomorrow, in a time preference task). In such instances, children can circle which of the options they prefer. However, very young children (below age 5) may still struggle with this task since their fine motor skills may not be sufficiently advanced. Hence, another alternative is electronic instructions (delivered on a computer or iPad) with pictorial representations that allow for clicking or tapping responses on a screen.²²

In sum, to maintain experimental control we suggest keeping language simple no matter the age of the respondents. Research shows that adults in America read at an 8th grade reading level on average (Wolf and Davis, 2004). Healthcare is one area where readability of materials is particularly important; hence, healthcare organizations provide guidelines for patient materials. A variety of national organizations have suggested that patient materials for adults be at or lower than a 5th, 6th or 8th grade reading level (see Badarudeen and Sabharwal, 2010 for more discussion).

3. *Executive functions:* Executive functions are higher-level cognitive abilities that encompass cognitive flexibility, attentional control, inhibitory control, and working memory (Naglieri and Goldstein, 2014). These abilities cut across many aspects of experimental design. Cognitive flexibility is the ability to switch between tasks or switch between different aspects of a problem. Attentional control is the ability to avoid distractions. Research suggests that these skills develop with age (Dajani and Uddin, 2015). We find it difficult to provide guidelines for how to design studies that incorporate this developing ability into experimental procedures. However, a useful rule of thumb is to avoid requiring children to “switch” their attention throughout the study:

²⁰To elicit risk and time preferences of their elementary school sample, Brocas et al. (2019) used electronic tablets programmed with Matlab. In this study, children viewed vignettes and made their choices by touching their preferred options on the screen.

this is why, for example, when we conduct experiments at CHECC, we are careful to use differently colored props for each decision task to avoid requiring children to work to switch their definition of a particular set of props (e.g., we use red/blue plates in a dictator game, and then use green/orange plates in a time preference elicitation task if it is in the same session.)

Working memory is the ability to consider and manipulate multiple facts simultaneously. Research on memory suggests that auditory memory is separate from visual and tactile memory, and that recall is better when information is presented via visuals (Gathercole et al., 2004; Cohen et al., 2009; Bigelow and Poremba, 2014). This suggests that experimental instructions should be presented visually when possible. Research shows that the main components of working memory are in place by age 6, and that their capacity increases linearly through early adolescence (Gathercole et al., 2004). Research requiring children to memorize numbers finds large increases in working memory with age, whereas research requiring children to memorize familiar objects documents a much more gradual development (for a summary, see Schneider, 2011). This follows because familiarity with numbers develops concurrently – children asked to memorize numbers perform poorly on the task if they are unfamiliar with them. The take-away here is that familiar objects and representations should be used when possible in economic experiments.

A related concept is focus and attention span. How long should experimental sessions reasonably last, based on the child's age? We were unable to find any direct guidance for this, partly because how attention is measured affects the results (in the parlance of hard scientists, the “Heisenberg Uncertainty Principle” at work, or in psychology the “observer effect” if the act of observation itself matters). Educators have developed rules of thumb based on some function of the child's age; e.g., multiply a child's age by 2, and this gives you the number of minutes that a child can pay attention. This implies that a 5-year old can usefully participate in a study for 10 minutes, a 10 year old for 20 minutes, and so on.²³ Until further research is conducted, we cautiously endorse this rule to maintain experimental control, though we suspect the rule should be non-linear, especially during the adolescent years.

Of course, the situation itself will matter as well, making it equally important to consider how to sustain attention. The literature reveals that children can pay attention longer if there are fewer distractions (Dixon et al., 2006; Fisher et al., 2014; Gaspelin et al., 2015). Hence, minimizing auditory and visual distractions is key for increasing the attention span of child participants.

²³ See, for example, <http://www.laurenqhill.com/understanding-the-of-attention-spans-of-elementary-aged-students/>

Another approach that works well is to take frequent breaks to engage in some physical activity, i.e., asking children to stand up and do a few jumping jacks (Ólafsson et al., 2013) and to keep children engaged by making sure that the instructions are interactive, i.e., by asking the child's input while proceeding through instructions (Morgan et al., 2002; Punch, 2002).

Inhibitory control is the ability to control impulses. The now well-known “marshmallow task” – where children who can avoid eating a marshmallow for some period of time receive two marshmallows - evaluates this skill (Mischel et al., 1970). More recent research in neuroscience suggests that inhibitory control is associated with future-oriented behavior in children (Steinbeis and Crone, 2016). Our research suggests that inhibitory control can be decoupled from time preferences by incorporating a front-end delay for rewards in the experiment (Castillo et al., 2020). More generally, experimental designs that trigger an impulse to make an immediate choice should be avoided, since we generally want subjects to make deliberative decisions that are not driven by self-control issues. This is a key assumption to maintain dominance in the experiment.

4. Abstract thought: Abstract thought is the ability to think beyond the concrete, for example, thinking about possibilities and forming new ideas. Abstract thought develops throughout the 3-17 year-old period and affects children's understanding of anonymous partners in a game or games involving possible outcomes that are difficult to represent with visual aids. Children begin to reason using abstract and symbolic thought as early as age 3-5, when they use ordinary objects in pretend play (Bergen, 2002). This suggests that children could understand anonymous partners at an early age. However, the abstract thought needed to represent concepts like probabilities is more challenging. For example, research shows that over- or under-estimation of probabilities can be due to concrete, rather than abstract, thinking (Lerner et al., 2016).

The ability to think about future events is part of abstract thought. Research suggests that 3 year-olds have trouble describing planned future activities, but that 5 year-olds are much better at this task (Busby and Suddendorf, 2005). In experiments eliciting time preferences, researchers have choices about how far in the future payments should extend. Some evidence suggests that children ages 3-5 perceive a 1-day delay as “long” (Schwarz et al., 1983). Hence, Andreoni et al. (2019b) use a 1-day delay for children ages 3-12. Most research with older children, however, uses longer delay periods on the order of 4-8 weeks (Bettinger and Slonim, 2007; Angerer et al., 2015b; Castillo et al, 2011).

5. Theory of Mind (ToM): The ability to separate one's own information set and desires from that of others - or ToM – is crucial for most social interactions and is relevant for many experiments we may consider conducting with children. Basic ToM usually develops around age 4. Prior to this age, children have difficulty understanding that someone else's viewpoint can differ from their own, which makes games with different information sets unreliable for this age group. By age 5, normally-developing children have developed ToM, though higher-order thinking skills continue to develop through young adulthood (Wellman et al., 2001), and within our CHECC data we find that, for 3-6 year olds, making the presence of other children salient increases their ToM (Charness et al., 2019). This is important experimentally since there is some evidence that children from lower SES families lag children from higher SES families (Devine and Hughes, 2018).

Evaluation of competencies

There are at least two reasons to evaluate child competencies. First, as discussed in Section 10.4, researchers may be concerned that concurrent development in cognitive abilities may create confounds and may therefore want to control for cognitive abilities. Standardized assessment tools are available for this purpose. Proprietary tests, such as the Woodcock-Johnson Test of Achievement (WJ) include sub-tests for numeracy (e.g., Applied Problems, Quantitative Concepts) and literacy (Letter-Word Identification, Spelling) that can be used with children ages 3 and up (Mather and Woodcock, 2001). The Peabody Picture Vocabulary Test (PPVT) is a good measure of receptive vocabulary (Dunn et al., 1965; Williams, 1999).

A highly recommended test for executive functions and other cognitive skills is the NIH Toolbox, which is administered on an iPad (Gershon et al., 2013; www.nihtoolbox.com). The NIH Toolbox includes sub-tests for cognitive flexibility, attentional control, working memory and language. Additional constructs include processing speed, immediate recall, and episodic memory. Most tests described above require 1:1 administration with a trained interviewer. If researchers have access to grades or standardized test scores, these can sometimes be used as proxies for cognition, since these are often correlated.²⁴

A second reason for measuring child competencies is that researchers may be interested in learning whether children can solve a specific problem or understand a specific concept in their

²⁴ Access to these materials has historically required providing evidence that the researcher has reasonable qualifications.

experiment. That is, they may wish to know whether behavior in their study is due to an inability to understand a task, or due to true underlying preferences. Likewise, such an exploration can yield insights on the cognitive effort children must make to complete the exercise. This understanding can help to ensure experimental control by providing an indication of the level of rewards one should use to ensure salience and dominance (Smith, 1982). This would not require a comprehensive assessment. Instead, researchers should develop their own questions. We discuss the use of such comprehension checks in Section 9.2.

9.2 Experimental Design by Child Development Age

Table 1 provides a summary that describes the range of methods that have been used in experiments with children and adolescents. The papers summarized in this table come from a comprehensive search of the literature using economic experiments with children, including papers from other fields that use incentivized economic games. The appendix describes our search process in greater detail. We identified 197 papers for inclusion in the table. Nearly half of the studies were conducted in Europe, about a third were conducted in North America, and most of the rest were conducted in Asia and Africa.

[Table 1 About Here]

We split the table's contents into five age categories: ages 3-5 (roughly translating to preschool age in the U.S.), ages 6-8 (early elementary school), ages 9-11 (late elementary school), ages 12-14 (early adolescence) and ages 15-17 (late adolescence). We do not include any studies of children below age 3 because there are very few (we identified only 7) and because we believe that children below age 3 generally do not have the competencies needed to participate in an economics experiment as we define them here. We do not include statistics from any studies of participants ages 18-22, since experimental economists have been conducting studies with university students for the past several decades. Instead, the last column of Table 1 provides the generally accepted standard for experiments with university students. Note that neuroscience research concludes that the brain does not fully mature until the mid-twenties (Pujol et al., 1993). How young adults differ from older adults is also an interesting research question but not one we address in this paper.

Below, we describe how features of experimental design in our review of the literature broadly change as we move from younger to older ages. We also compare the methods used with

children to methods typically used in laboratory experiments with adults. This provides a summary rather than a set of recommendations, though as we outline below, many times the commonly used method in each age category is also most appropriate given the child competencies discussed in Section 9.1.

Instruction Delivery Format: This refers to the number of children participating per experimenter (one-on-one or in a group). 55% of experiments with 3-5 year-old children use one-on-one implementation, while this changes to small group and to larger groups with age (e.g., 66% of 15-17 year-olds are in a group). By comparison, laboratory experiments with university students include a typical session size of 20-30 students, limited by the number of computers in a standard lab. Given the limited ability of children to focus attention and ease of being distracted discussed in Section 8.1, as aforementioned, we believe that one-on-one administration for younger children is appropriate. Yet, this approach does not come for free. Such one-on-one administration leads to the assessor potentially playing a much larger role in the experiment. This is because with young children, instructions are spoken instead of read, responses are usually verbalized instead of written, and with the one-on-one temporal constraints, if there is variation in assessor quality, different subjects are receiving different versions of the treatment, violating the Stable Unit Treatment Value Assumption (SUTVA). We return to this issue in Section 10.

Identity of partner: The identity of the partner can be relevant, especially for young children who may not have sufficient skills to employ abstract thinking skills. Hence, for children ages 3-5, it is common to see the partner in person (33%) and relatively less common for the partner to be anonymous (20%). Alternatively, for later ages it is most common to have an anonymous partner (about 50% of the time for ages 12-17) and indeed this is also the norm for standard laboratory experiments using undergraduates as subjects.

Instructions and Instruction Aids: This category refers to how instructions are presented to participants. Instructions are read aloud most (70-90%) of the time for all ages, and this is also true for standard laboratory experiments, where instructions are usually read aloud to a group while participants follow along. Physical props-e.g., plates/envelopes as containers of items for recipients, or physical tokens to be allocated among recipients - accompany instructions about a third of the time among ages 3-5 and 6-8 and about a fifth of the time for ages 9-11. Physical props are rarely used for adolescents, and rarely used in laboratory experiments with adults (with the exception of bingo cages to reveal payouts used in some lab games).

Graphical aids are used in 15-18% of the cases for ages 3-11, and more frequently - about 25% of the time - for children ages 12 and up. Such aids are useful to harness foundational math knowledge discussed in Section 9.1. Electronic instruction delivery is rarely used, but when used it is a more common format for older children who are more capable of using these devices independently. Overall, electronic devices are utilized roughly 5% of the time for children below age 14 and used in 15% of the experiments for children ages 15-17. Since electronic devices have become cheaper and easier to program over time, the rate of their usage has increased significantly since 2015. In experiments from 2015 onwards, electronic instructions were used in 80% of the experiments for ages 15-17.

Subject decision format: Subjects can make decisions in an experiment by pointing or moving objects (physical), verbally (oral), in writing by circling a response or writing in a response (written) or by clicking a button on a computer (electronic). About two thirds of 3-5 year old experiments feature physical subject decision format, which is important here because as we discussed in Section 9.1, receptive vocabulary emerges before productive vocabulary. Moreover, a written format is never used until subjects are at least 6 years old, and even in that case it is used in only about 25% of experiments. This is also reasonable given that writing abilities are only emerging at early elementary age. By comparison, for adolescent subjects ages 13-17, decision format is more likely to be written (about 50%) or electronic (about 25%). This mirrors closely the standards for laboratory experiments, which are either conducted using paper-and-pencil with written responses, or conducted electronically with ZTree (Fischbacher, 2007) or similar software.

Comprehension check: We identify whether there was a comprehension check before proceeding to the decision, which could be oral, physical, written, or electronic. Comprehension checks are often used in standard laboratory experiments and should be even more important for young children given the uncertainty about what they are able to understand. Surprisingly very few studies across age groups mention comprehension checks – comprehension checks are documented in only 30-60% of experiments across all age groups, with seemingly no pattern by age. A potential reason for this result is that comprehension checks are used, but not mentioned in the description of the experiment; or that given the shorter session times, are not used to save time. The comprehension check strikes us as a valuable addition to any experiment with children – especially given the limited cognitive abilities of children relative to adults - and we therefore

suggest that experimenters use this tool and report the results of comprehension checks in their papers.

Total number of subject decisions and experiment length: This category refers both to the time subjects spend in the experiment (in minutes, with minimum, maximum, mean, and standard deviation), and the number of decisions that subjects make in the experiment (again, with minimum, maximum, mean and standard deviation). Among children ages 3-11, subjects make 6 decisions on average. This increases to 13 decisions for ages 12-17. Experiments take 35 minutes on average for children ages 3-8 and 40-50 minutes on average for children ages 9-17, all of which exceed our recommendations on experimental duration discussed above. By comparison, laboratory experiments with adults tend to last 45 minutes to 2 hours.

Payout for repeated decisions: In cases where subjects make many decisions, it is common in standard laboratory experiments to pay out only one randomly selected decision to avoid confounds caused by paying out all decisions, such as income effects and portfolio effects. But children may have difficulty understanding the concept that only one decision is paid. Indeed, over 80% of experiments in the 3-8 year old age group pay for play in all rounds. We find that it is more common to pay only one round at older ages, with nearly 50% of experiments paying for choices in only one round at ages 9-17. Paying for choices in one round, while a common strategy in typical laboratory experiments, is less common in experiments with children. This is reasonable, since as we noted in Section 9.1, children have difficulty understanding probabilities, and paying out one round adds an additional layer of complexity, in some cases likely causing dominance or salience to be violated.

Incentives: We identified four main approaches for incentivizing children in the literature: giving out a pre-set prize such as candy, a sticker or small toy, giving out tokens that can be exchanged for prizes in a “store”, using gift cards, and money. Our survey of the literature finds that common prizes include packets of Skittles, crayons, and balloons for ages 3-11 and chocolate bars, pens, and markers for ages 12-17. The appendix provides a detailed list of the prizes used by age group (see Table A1).

As shown in Table 1, for children ages 3-8, a prize store is the most common incentive (about 35%), and candy is the second most common (about 20%). The store is also the most popular at ages 9-11 and is used nearly in half of the studies. The second most popular approach at this age is to use money (about 25%). After age 11, money becomes more popular and is used

in 54% of the cases for ages 12-14, and 76% of the studies for ages 13-17. Gift cards are less common and are used in only 6% of the studies for ages 13-17. By comparison, the standard in laboratory experiments is to use money. In our view, this difference in payoffs over the age range is reasonable since the payment medium should yield a salient reward where the subject's utility is increasing in the reward and non-satiation holds. Of course, incentives are a cornerstone of economics and therefore deserve more careful attention. In Section 10.7, we highlight important considerations when using incentives with children.

Exchange rate: When an exchange rate from points earned in the experiment to prizes or money is used, it is most common for this exchange rate to be 1:1 at age 3-5 (55%) and to be more complex for ages 6 and up (50-70% of the time, increasing with age range). Exchange rates in laboratory experiments also tend to be more complex than exchanges in experiments with children.

Types of Experimental Games: The last row of Table 1 provides information about the types of games used in the experiments that we surveyed. A large proportion of economics experiments used with young children ages 3-5 were dictator games (46%), and this is partly because this is one of the easiest games to explain and partly because of the rise in popularity of dictator games among developmental psychologists, who are interested in the evolution of social preferences at a young age. Starting at about age 9, many standard economics games become well-represented (e.g., games eliciting time, risk and competitive preferences; games of strategy; ultimatum games).

Comparing Methods: A limited set of papers investigate how design features affect children's decisions. Bunch et al. (2007) test a variation of the children's gambling task, wherein they introduce either a "gains only" or a "losses only" deck against a variable deck. With this variation, they lower the relational complexity and find that all children succeeded in these tasks. Angerer et al. (2015b) compare two methods of time preference elicitation in a large group of children, finding that a multiple price list and a single choice measure yield similar results in the aggregate. We urge future research to explore such methodological considerations, as a maturation of this field will depend on such new insights.

With the extant literature and its many heterogeneities in mind, we turn to a set of ten tips that serve as guidance for the researcher interested in conducting field experiments with children. The ten tips might also help consumers of this literature more deeply understand the various design choices and their consequences.

10. Ten Tips for Pulling Off Experiments with Children

Sprinkled throughout the previous sections are numerous experiments conducted with children as subjects. The experiments sometimes leverage randomization for identification of the key treatment insight, sometimes leverage natural variation by judiciously applying experimental control in that setting to obtain preferences or other information to inform a model, and sometimes do both. Regardless of approach, the literature teaches us that similar to the spirit in which astronomy draws on the insights from particle physics and classical mechanics to make sharper inference, experiments with children can provide principles to sharpen inference around economic issues of import. Likewise, such experiments can help to determine whether received insights should be reinterpreted or defined more narrowly than first believed.

Our own experiences with experiments with children range from baseball card shows in the early 1990s to open air markets in the early 2000s to starting our own schools in Chicago and Bangladesh to lab experiments in several parts of the U.S. We have learned that the notion of conducting controlled studies with children might appear daunting. Figuring out the market, how to randomize relevant features within a complex situation, and avoiding pitfalls is not taught in standard economic or even social science courses.

In this section we offer ten tips for executing successful experiments with children. We suggest that all economists conducting experiments with children carefully consider these tips, which range from tips about idea generation, to avoiding pitfalls that are present in much of the current literature, and practical considerations such as recruitment and incentives. These tips were borne out of our interpretation of where the literature with children should be taking us, as well as discussions with colleagues who had practical questions about how to conduct experiments with children. These tips are not intended to be treated as universal truths; rather, they are intended to help researchers think through the design and implementation of experiments with children.

The first two tips drive home the point that it is important to consider theory and mechanisms when developing research agendas using children. The next two tips remind researchers to be cognizant that children are not just “little adults.” Children develop rapidly in childhood and adolescence. It is critical to incorporate measurement of this development and to acknowledge potential limitations in interpreting findings due to this development. The fifth tip straddles interpretation issues and practical design issues. It focuses on recruitment, and, in particular, reminds researchers to seek representative samples or to be cautious of drawing broad

conclusions from non-representative samples. The next four tips provide practical advice related to procedures, incentives, ethics concerns and documentation. Finally, the last tip discusses scaling concerns, which is relevant for researchers designing interventions for children.

10.1 Use economic theory to guide your design and as a lens to interpret your data

This tip is ubiquitous, and any researcher interested in experimentation should begin here. Indeed, this is tip #1 from List's (2011) 14 tips for pulling off a controlled field experiment. Yet, it bears repeating here since economic theory is portable. By themselves, data from experiments with children can only offer a limited understanding about what is likely to happen in a new setting or with a new population of children. Incorporating guiding principles or hypotheses that drive the research agenda – whether this is a formal theory or an informal conceptual framework – is key for guiding the interpretation of results. Together, theory and experimental results provide a powerful tandem to learn about the world and provide prescriptions for improvement.

10.2 Go Beyond A/B Testing to Deepen our Understanding of Children

A natural starting point from 10.1 is to explore mediators to observed relationships. Such mechanism exploration begins with theory, which is a key guide determining a path deeper than simple A/B testing. There is a large literature in human capital development that relies on A/B testing – i.e., comparing educational interventions to control groups – to draw inferences relevant for public policy (e.g., Heckman et al., 2010; Attanasio et al., 2012; Sylvia et al., 2020). This is an important approach that highlights the value of experiments with children. Throughout this article, we have also highlighted how we can use experiments with children to deepen our understanding of the underlying mechanisms behind the data patterns observed. In this manner, in addition to measuring the intervention effects (e.g., does this new Kindergarten curriculum lead to higher test scores?), researchers should design subtreatments tightly linked to economic theory to reap the true benefits of the experimental method by exploring why it works (e.g., a sub-treatment, or information gathered in a complementary survey, which shows the channels through which the Kindergarten curriculum worked). Getting to the whys of an intervention also have important effects on scaling, as discussed in Section 10.10. Czibor et al. (2019) provide a deeper discussion of going beyond A/B testing in broader economic settings.

In a nutshell, the combination of 10.1 and 10.2 provides a natural umbrella for a research agenda with children. As an example, if the researcher is interested in human capital formation, one first examines the optimal timing and type of investment over the life-cycle, such as characterizing the types of interventions that are optimal. At that point, explorations going beyond A/B tests determine the important moderators and mediators of the examined interventions (see, e.g., the work at CHECC cited above using traditional experimental approaches and the work of Ye et al. (2021) for an example using EEG). With that knowledge in hand—the why’s, where’s, when’s, and for whom—applicable theories and optimal policies are better understood, and the basics for scaling are in place.

10.3 Align Child Competencies to Experimental Design

We next turn to issues that are specifically relevant to studies of children. The importance of aligning competencies to experimental design cannot be over-emphasized, since children are not simply “little adults,” and the developmental process is complex. This underlies the goals of Section 9, which is devoted to describing how child competencies develop with age and the key design features that should be taken into account when designing experiments with children and adolescents. Aligning child competencies to experimental design requires an understanding of cognitive skills and executive functions, and Table 1 in Section 9, which summarizes how related work has tackled this issue, can help guide the design of the experiment.

When studying how child preferences develop with age, competencies that develop around the same age range can be a confound. This is especially true for cognitive abilities. For example, research in developmental psychology shows that a young child’s ability to split rewards equally is associated with math ability (Chernyak et al., 2016). Research in economics also shows correlations of cognitive abilities with child preferences (Benjamin et al., 2013; Andreoni et al., 2019a), and suggests there is an association between theory of mind development and sharing behavior (Cowell et al., 2015). Competency development can also be related to SES since children from low-income households tend to have lower literacy and numeracy skills than their higher-income peers at the same ages (Liaw and Brooks-Gunn, 1994; Noble et al., 2007). Moreover, consistent with our framework outlined in Section 4, within the same age range and background, children can vary greatly in their level of cognitive development (Sternberg, 2011). Hence, it is important to consider the associations between competencies and decision-making in the

experiment. In the next tip, we discuss assessment of competencies as an important way to avoid confounds. Importantly, experiments that are not designed to be appropriate for the range of cognitive abilities in the study sample may introduce serious endogeneity issues if subjects who do not understand the experiment cannot participate due to lack of understanding.

As aforementioned, when we align child competencies with optimal design, we might be presented with unique problems. One such example is if one-on-one administration is necessary (as we suggest in Section 9 for 3-5 year-olds), and there is variation in assessor quality. In this case, a key assumption underlying the internal validity of the experiment is violated: the Stable Unit Treatment Value Assumption (SUTVA). This is because different subjects are receiving different versions of the treatment, rendering any simple comparison between treatment and control groups uninterpretable. We therefore urge the researcher to devote special attention to how assessors are allocated to subjects, making sure that assessors are distributed evenly across treatments, and examining if the assessor fixed effects are important to outcomes.

10.4 Consider Stability and Confounds with Cognitive Development

The timing of preference development is an important question. For example, when applying theories of rational decision-making to children, it is necessary to consider at what ages children can choose rationally. When designing interventions for children, one needs to have some idea of utility functions that would allow policymakers to assure that interventions are appropriate (e.g., providing immediate incentives for younger children, who are more impatient than older children). Finally, learning about how preference development is assisted by or interacts with cognition can give us a deeper understanding of what key milestones are needed for economic preferences to develop.

When evaluating the development of preferences with age, it is important to consider stability of measurement. For example, if economic preferences are unstable with age, does this mean that the underlying processes affecting this development are not stable or that the elicitation methods used are not stable? One way to evaluate stability while reducing measurement error is to elicit preferences using several methods (as discussed in, e.g., Gillen et al., 2019; Angerer et al., 2015b used two methods to elicit preferences, as do Datar et al., 2019; Samek et al., 2019) and account for decision error (e.g. Castillo et al., 2018). When eliciting preferences over a large age range, different methods were often used with different age groups (e.g., as in Andreoni et al.,

2019b). When using different methods for different ages, it is useful to have overlap in the methods and age ranges to assure that there is a correlation with methods.

Cognition itself can also be an issue since it changes over time. Hence, if economic preferences are correlated with cognitive ability, but cognitive ability is not measured, then it is difficult to determine whether cognition is one channel through which age-related development affects preferences. Further, if researchers report that time preferences are correlated with a particular field outcome, it is unclear whether time preferences themselves, or a correlated factor, drive that relationship. Measuring and controlling for cognitive ability is one way to improve on existing methods; and this was done in several CHECC papers (e.g., Cappelen et al., 2020; Castillo et al., 2020). For example, Castillo et al. (2019a) report that time preferences measured early in life are associated with later field outcomes, even while controlling for SES, cognitive abilities, non-cognitive abilities, and other economic preferences.

Measuring cognitive ability also helps with another problem, which is that lack of sufficient cognitive skills to understand a decision task can result in greater errors in the task at younger ages than at older ages, reducing the precision of the estimates at younger ages. One way to assess whether difficulties in understanding the task affect the results is to include questions on understanding in all experimental instructions (and to record the child's answers to these questions). Relatedly, insufficient non-cognitive abilities could affect responses to a task. For example, Brocas and Carillo (2018) found that children may have the ability to think strategically but may not have sufficient noncognitive skills to act strategically.

Children who do not understand the experiment may either not respond or may provide random answers. Researchers may need to impute data or drop data in this case (dropping data seems to be most common). Such treatment of the data has a major limitation, which is that the subjects that remain would be a selected (higher cognitive ability) sample, which will introduce endogeneity and affect interpretation. A far better alternative is to create an experimental design that is suitable and easy to understand for the lowest-performing children in the sample. This can be achieved through following our guidelines in the prior tip and doing iterative pilot testing.

An area that has received less attention in economics than it has in psychology, psychometrics or medicine is test-retest reliability (DeVon et al., 2007). Test-retest reliability is relevant for constructs that are not expected to change over time. Conducting experiments with the same group of children over a short time period would help to illuminate exactly how much error

exists in the measures that we use (see Gillen et al, 2019, for such an approach with adults). Designing batteries of tests all aimed at measuring a particular economic preference would help to reduce this error problem.

10.5 Recruit a Representative Population

This tip straddles issues of interpretation and practical concerns. With regards to interpretation, if the researcher wishes to make broad conclusions about children, it is important to either recruit a representative sample or to acknowledge the limitations in terms of drawing broad conclusions from a sample that is not representative. Most studies with children use selected samples. Unlike standard laboratory experiments with university students (also a selected sample), there are several ways in which subjects can be recruited, and in turn, several ways in which child samples can be selected. Selection can occur on the basis of the SES of the participating school, buy-in from only a sub-set of parents or even based on criteria for cognitive abilities required to complete the task. Depending on the research question, selection may be desired. For instance, a researcher wishing to study economic preferences of children with specific mental illness would wish to target recruitment to a sample of children that is representative of the population of children with mental illness. Below, we outline several popular approaches for recruiting children and discuss the benefits and drawbacks of each.

Recruitment of child populations is the first obstacle in conducting research with children. A popular approach in economics is to forge partnerships with schools and contact parents through these schools to invite participation (see, e.g., Angerer et al., 2015a; 2015b; List and Samek, 2015; Castillo et al, 2011; Houser et al, 2015; Harbaugh et al, 2001). A major benefit of such an approach is that schools have large lists of eligible children, and parents may be more likely to participate if they receive the information from a trusted source. Further, schools can often benefit from the knowledge gained from the study. However, a potential drawback is that many schools are already overwhelmed with research requests and might not have the capacity to engage in another project. Also, some districts have additional application processes before allowing research with their students. If parent consent is necessary for the research, researchers should use a multi-pronged approach to engage parents, including sending flyers home in backpacks, as well as recruitment events on site and phone calls. This is because many parents will not return a consent form sent home in a backpack, but they may respond when provided with more information. If a school

partnership is possible, we believe this is one of the best ways to proceed due to the large potential sample size available and the collaboration between researcher and schools. Schools are also a good idea if researchers wish to follow-up with children at a later date, since schools could provide updated contact information for children if they have not moved out of the school district. Another advantage is the ability to link children with administrative data (e.g., grades, standardized test scores or disciplinary referrals) with parent and student consent.

Some researchers cite difficulty in getting school buy-in as a major obstacle to recruitment and propose alternatives. An alternative approach that has been gaining popularity among developmental psychologists is conducting recruitment in public places, such as museums and public parks (see, for example, Blake and McAuliffe, 2011). This includes setting up a table at a location and inviting families who pass by to participate. Museums sometimes charge researchers to use the space, while public parks are free (though may still require permission from the city). This method can be effective if the researchers can provide credibility, parents are not rushed, and their child can participate on the spot. But the drawback for researchers wishing to recruit subjects for an intervention or longitudinal study is that it could be more difficult to keep track of populations recruited this way. Another drawback is that if the research question involves a particular sub-population of children (e.g., those with mental illness) then attempting to recruit in public spaces may be time consuming and sub-optimal.

Another possibility is recruiting through existing parent networks such as Parent-Teacher Associations (PTAs) affiliated with nearby schools and asking them to forward information about the study to parents (Foss et al., 2010 cites this as one of their most successful approaches). This may work better than recruiting at museums or public parks if researchers are seeking children with a particular characteristic (e.g., children of a specific age). Another alternative along these lines is to post messages in social media groups, which has the benefit of being able to select groups associated with a particular characteristic that is targeted (e.g., mental illness). This approach does require interested parents to follow through with bringing their child to the experiment.

Finally, one can obtain registry data to mail invitation letters to families. For example, county birth records are used to send mailed recruitment letters to parents of infants for some psychology studies (e.g., Kim and Johnson, 2013). Registry data is also used in Deckers et al. (2015), who saw a recruitment rate of about 12% from their mailings (they offered participation

in a free program as part of the study incentive). While we have been unable to find recruitment rates for related psychology studies, we suspect it to be substantially lower as many of these studies do not offer incentives.

Having a representative sample of children in the study is important for drawing general conclusions or making claims of external validity. The registry data approach may give researchers the highest chance of obtaining a representative sample since it involves recruiting a wide sample of the population, though even then the participants who choose to participate are likely to be selected. Focusing the study on one school reduces representativeness since children within a single school are likely to be representative of only a single area, hence an alternative to improve representativeness is to select schools with children from a range of backgrounds.²⁵

Suggestions for reaching out to parents include using simple language to communicate study goals (this can be very broad, e.g., understanding how children develop) and making study participation as convenient as possible (i.e., nearby study locations, catching the parent at a time they are not rushed and offering participation on the spot, providing flexible scheduling options). Offering an appropriate participation payment and gift for the child can also incentivize parents to participate.

10.6 Adopt Appropriate Study and Experimental Procedures

This tip and the following three tips deal primarily with practical concerns. An important consideration is where and how the research should be conducted. A standard approach in developmental psychology for research on children is to ask parents to bring the child to a child-friendly ‘laboratory’ on campus. Such a facility should include age-appropriate seating, provisions for a parent waiting area, and adequate parking. Economists typically will not have a facility like this at their disposal. But, the good news is that economic experiments do not require major equipment, so setting up a “mobile” laboratory is not difficult.

One possibility is to develop a partnership with a school, which is already child-friendly, and then either conducting experiments during class time or by asking parents to accompany children to the school in the afternoons or weekends. Another alternative is to recruit children to participate at the recruitment location, for instance at public parks or children’s museums (more

²⁵ An example of such an approach is Brocas and Carrillo (2019), who conduct their study in a private school in a relatively rich area and a public school in a relatively disadvantaged area.

on this in the next tip). A major benefit of these alternatives relative to participation on a university campus is that it does not require the parents to travel very far to participate. However, a major drawback is that not all schools will be amenable to this approach – for example, poorly funded schools in low-income developing countries may be noisy, disorganized, and violent. Another drawback is that even if the school is conducive to research, more extensive set-up time will be required at the field sites; for example, it is a good idea to try to remove visual distractions by putting up cardboard dividers and to look for areas that do not have high distracting noise levels.

Another possibility is to conduct experiments in participants' homes. In this approach, the experimenters (teams of 2 are a good idea for safety reasons) arrange and then conduct home visits. This approach is also not without drawbacks, since particularly in low-SES environments, homes may be overcrowded and full of distractions. Further, not all families will feel comfortable with researchers visiting their homes.

Since many of the experiments we describe (especially those with young children) rely on verbal and physical interactions, the capabilities of the experimenter are of vital importance. Experimenters should be experienced in interacting with young children. A number of training sessions specific to the protocol should be conducted to assure that the experimenter has memorized the instructions and can easily carry out the experiment. In our work, we often train undergraduate students at the university to conduct sessions. The university students we hire most often come from fields such as public policy, social work, and education.

The above has provided several alternatives for where to conduct experiments with children. We are not aware of data that explores the association of study environment with outcomes, but we think it is reasonable to assume that reducing the level of distraction in the environment will lead to higher precision of estimates.

10.7 Understand How to Use Incentives to Motivate Children

The topic of incentives arises most frequently when discussing design features of experiments with children. The pioneers of the experimental approach argued that incentives are an important component of the experimental environment because of the control that they provide. In particular, the reward medium must satisfy three conditions: non-satiation, salience, and dominance (Smith, 1982). As discussed earlier, cash incentives, which are typically used in laboratory experiments with university students, may not work for all age groups. Table 1 presents

a progression of incentives by age, which shows that incentives at early ages most often take the form of candy or stickers, middle-childhood incentives often take the form of prizes, and adolescents can be compensated with gift cards or cash. This progression is appropriate, since children are not usually shopping on their own until middle to late adolescence.²⁶ When considering non-cash rewards, many begin with attempting to solve heterogeneity issues. While heterogeneity over the valuation of the reward medium does not present problems of internal validity if the groups are balanced across treatment and control, the use of non-monetary rewards does present unique challenges in other ways.

For example, consider the case of providing non-cash rewards, such as stickers, to 5-year olds and asking them to play a dictator game. This is usually done to measure social preferences among young children. The researcher wishes to interpret “number of stickers shared” as a social preference measure, yet such a measure is easily confounded with, e.g., the child’s preference for the certain types of stickers in his/her endowment, by the child’s preference for stickers more generally, or by the child’s expectation about the preferences toward stickers of his or her partner. While these issues are also present when considering the allocation of dollar bills of University students in standard dictator games (utility value of money is indeed heterogeneous), dollars can buy a myriad of goods and serve a useful intermediary role in modern economies. Alternatively, there is not a ready resale market for stickers, and therefore they are best considered for final consumption. As such, in experiments of this ilk, we learn a lot about preferences over certain stickers, but perhaps nothing more generally without further assumptions. Alternatively, when we use a medium of exchange, such as currency, as an incentive, we learn about preferences over that medium, and non-satiation, salience, and dominance are more likely to hold.

One way to begin to address this issue is by allowing the child to first choose stickers that he or she most prefers to play with, as in Cox et al. (2016). Similarly, children may be asked to sort potential prizes in order of desirability, and the top few are chosen as the incentives for future games (Rice and Broome, 2004). This approach assures that the children value the incentive. Another possibility is to use a prize store. Here, the children make decisions with tokens that are exchanged for goods in the prize store. Provided the prize store is stocked with goods that all children like, and they assume other children like them as well, then this is closer to using monetary

²⁶ Note that some local IRBs may not allow incentives, and that schools or other spaces where the experiment is conducted may have guidelines or limitations on types of incentives used.

payoffs, although the payoffs are made in lumpy, discrete amounts. There is a problem of potentially large diminishing marginal utility over two of a similar item; therefore the prize store should have a large enough variety to limit satiation.

The prize store is flexible, since researchers can include prizes, toys, school supplies and so on that appeal to different age groups and tastes. The prize store is also easily explained – for very young children, a 1:1 exchange rate of points earned in the experiment to prizes can be used, which limits the need for calculating prices of items. Using a prize store also allows for comparisons across age groups or for longitudinal data collection, since the store can be set up with items that appeal to all age groups. For example, in studies which children and their parents, children could earn points to exchange for small toys, while parents could earn points to exchange for small items like snacks, office supplies or travel toiletries.

The other aspect that frustrates interpretation in experiments with children is that children are often rewarded by parents or teachers for doing what the adults want—and punished for doing what the adults do not want. Therefore, the potential for demand effects – especially when choices are not made privately – are potentially large. In addition, parents may expect children to share prizes with siblings after the experiment. Hence, measured preferences can be confounded with expectations about background consumption. The former issue can be minimized to some degree by allowing for private decision-making whenever possible. We have previously attempted to minimize the latter issue by including notes sent home to parents asking them not to require their children to share their incentive (although, such notes are more useful in repeated experiments, since in the first experiment children may expect a sharing requirement).

The problems raised in this section are more relevant for studies attempting to measure economic preferences than to studies considering treatment effects. Interpreting the level of an economic preference – e.g., social preferences, risk preferences, or discount rates – relies on making assumptions about valuation of prizes. The issues are less problematic in experiments that seek to measure treatment effects if an assumption is made that valuation for incentives does not interact with treatments, provided salience, dominance, and non-satiation are satisfied.

In summary, we recommend prize stores be used for very young children whenever possible. Children begin having more independence over spending their own money to purchase snacks or toys by around age 9. If the researchers do not intend to compare to a younger age group, then money should be used from this point forward. Money could take the form of gift cards if the

experiment is conducted in schools that prefer not to distribute cash or if the experiment is in the field and the experimenter(s) prefer to limit the amount of cash they are carrying at one time. We found that the most commonly used type of gift card varies by country. Walmart and Amazon gift cards were commonly used in the U.S. (Castillo et al., 2011; 2018; Barash et al., 2019), while Bol.com vouchers were used in the Netherlands (Bekkers, 2017).

10.8 Obtain IRB approval and Pre-Register Experiment

This tip revolves around two key pre-experimental administrative tasks that are invaluable. Indeed, this tip applies to experiments with children, but also spans all work that involves human subjects. We will discuss each in turn. In the United States, and many other developed countries, the Institutional Review Board (IRB) or ethics committee provides oversight into research activities involving human subjects. Some US-based researchers may wish to conduct human subjects work in developing countries. They may therefore face additional regulations based on local laws and IRB requirements. Whether researchers have obtained local IRB and even if there is no local IRB, researchers must still obtain IRB approval from their home institution. Much of the below provides guidance for the typical rules at US-based IRBs. Researchers based in or working outside of the US should do their own investigation to determine what rules apply.

Children are considered a vulnerable population and require special additional protections. However, most IRBs are open to experiments with children, provided that the researchers have identified a clear rationale for why children are included in the study and have addressed ethics concerns specific to children and broader ethics concerns that would also apply to studies with adults. IRBs are familiar with research with children since many researchers in departments such as psychology and education conduct research with children.

Most studies described here will generally fall into the category of minimal risk. As such, the only additional consideration is that the consent process proceeds in stages. First, the parent will be asked to complete a consent form on behalf of their child. Usually, the consent of one parent is sufficient.²⁷ Second, children will be asked to agree to participate (called “assent”) either verbally or on a written form. Each IRB may have their own requirements for assenting children,

²⁷ As noted by a reviewer, some areas where research might be conducted may have low literacy rates or no written language. In these cases, written consent may not be practical, and the researcher should work with his/her IRB to develop appropriate consenting procedures.

and they will usually consider the age and maturity of the child to determine a) whether assent is required and b) whether assent needs to be documented in writing or can be elicited verbally.

Studies with children that provide an intervention or require longer-term follow-up have a more in-depth IRB process than one-off studies measuring an economic preference. For example, because the CHECC program featured an early childhood intervention administered by the researchers, as well as a consent form allowing for 14+ years of long-term follow-up, the CHECC study underwent full board review. This included a presentation to the IRB regarding the nature of the study and the plans to minimize risks.

Just as with studies that use adult subjects, a key ethical consideration is that children do not feel coerced into participation – either from experimenters, teachers, parents, or peers. Ways to minimize potential for coercion from teachers or peers are to have children choose whether to assent in private. Ways to minimize potential for coercion from parents are to stress to parents that it is up to the child whether or not he/she wants to assent and to not withdraw the parents' participation payment if the child does not wish to participate.²⁸ And, incentive payments to children – which often take the form of gift cards or small gifts or candies – should not be so extreme that children feel compelled to participate or to share their payment with parents, siblings or friends.

In many cases, obtaining consent from the parents of all the children that are eligible for the study is challenging. For example, a researcher may be conducting a school-based intervention, but only some parents have consented for their child to participate. The most typical solution is to offer the program to all children in the study, but only evaluate or keep the data from those children who have consent on file. In other cases, for example if the research involves minimal risk and would be undertaken during the school day anyway, the IRB may allow the researchers to waive consent and instead provide parents with information sheets with the ability to opt out. In our experience, the proportion of consent forms that are returned is highly dependent on the SES of the school as well as the buy-in from the teacher (who can nudge parents to return forms).

Section 10.5 discussed various ways to recruit participants. All of the proposed channels should be documented for the IRB to assure that they follow guidelines and are not considered

²⁸ There are also possible costs to paying parents based on their child's performance in a task, and such interventions should only be considered after careful ethical review. For example, parents who do not receive an expected reward due to their child's performance may retaliate against their child, which would be a concern to consider carefully.

coercive. For example, contacting parents or students through social media would require IRB approval. And, it is against the policy of most IRBs that we are aware of to provide referral bonuses or incentives that are linked to recruitment of subjects.

Finally, just as with adult subjects, a careful consideration of costs and benefits should be undertaken. Children should not be used as a convenience sample – that is, just because one can go into a school to conduct experiments with children, does not mean that they should be used instead of adults. Child subjects should be used when the research question would benefit from the participation of children in the study, such as when the researcher is studying the development of economic preferences in childhood.

While IRB approval has been around for decades, a relatively new (at least for economics) pre-experiment task of import is pre-registration. Established in May 2013, the AEA-RCT registry provides a venue for researchers to document their experiments in a manner that is searchable by external audiences (Abrams et al., 2020). Researchers document the specific details of the experimental design, target number of observations, etc. and can also include a pre-analysis plan detailing how the researcher plans to analyze the data.²⁹ In theory, if used appropriately, this innovation can tackle key issues in the credibility crisis, most specifically, the “file drawer” problem and “p-hacking.” Both issues are key to solving the false positive problem. As Abrams et al. (2020; p. 1) discuss, the registry thus far has had quite limited success: “over 90% of randomized control trials (RCTs) in economics do not register, only 50% of the RCTs that register do so before the intervention begins, and the majority of these preregistrations are not detailed enough to significantly aid inference.” While other registries show similar impotence, and therefore provide limited advice on improvement, as part of this tip we follow Abrams et al. (2020) and urge experimentalists to at least submit their IRB materials in the pre-registration *before* any collection of data.

10.9 Document Procedures and Promote Replication

Beyond attenuating false positives via pre-registration, replication represents a key feature that can lower long-run false positives. Proper documentation of experiment procedures is key for

²⁹ A researcher can register a study without submitting a pre-analysis plan and can submit a pre-analysis plan without registering (although not on the AEA-RCT site).

future replication.³⁰ Laboratory experiments have benefitted from unwritten standards about what is important to report – and researchers who report on laboratory experiments typically provide information on session size, experiment length, average payout, how instructions were provided, whether privacy dividers were in place and whether questions on understanding were used. Further, a standard in laboratory experiments is to provide experiment instructions in an appendix.

We propose that similar standards be adopted for experiments with children, since many papers that we surveyed did not report on these key factors. For example, some papers failed to report on the session size, the exchange rate between tokens received in the experiment to prizes in the prize store, and a description of the setting, e.g., level of privacy and presence or absence of distractions. Papers should also report on the environment in which the study was conducted – including details such as level of privacy and distraction – which could help researchers understand how important such details are for the measurement of outcomes. These variables are all important for future attempts at replication and for moving the methodological literature forward.

10.10 Use Backward Induction to Design Interventions that will Scale

A complement to the representativeness of the population discussion in Section 10.5 is the representativeness of the situation. Indeed, to take both discussions to a higher level is to consider whether your insights will scale. This tip is aimed at the first silo that deals with conducting experimental interventions. When doing that type of experimental intervention, researchers should always be aware of how our interventions scale or have effects at levels beyond the setting of the experiment. To be more perspicuous, consider the “voltage effect,” from the implementation science literature. When considering the effects of programs at scale, a common finding is that the effect sizes diminish substantially (see Al-Ubaydli et al., 2017a, 2017b). We are only at the beginning of understanding the science of using science, but there is some evidence that shows there are good economic reasons why program effects fade at scale, and how benefit/cost ratios change at scale. Al-Ubaydli et al. (2020) examine the market for knowledge and show that the incentives of the various actors (researchers, government policymakers, citizens) themselves can lead to the scale-up problem. Importantly, if researchers backward induct when setting up their original research plans to ensure accurate and swift transference of programs to scale, not only are

³⁰ Further, proper documentation is also important for pre-registration and pre-analysis plans, which are becoming necessary standards and conditions for publication.

initial insights useful, but they aid in the generality (external validity) problem. Al-Ubaydli et al. (2021) provide a checklist for both researchers and policymakers that involves design choice from idea inception to roll out at scale.³¹

11. Discussion and Conclusions

Experiments with children are uniquely situated to provide insights to economists and policymakers. Our framework proposes that child outcomes depend both on genetics and environment, and that investments by parents, schools, and the children themselves affect these outcomes. The literature that we survey documents large heterogeneity in preferences even at an early age and finds evidence that parenting and school-based interventions have an impact both for the formation of child skills and economic preferences. Our model further suggests that child economic preferences affect both contemporaneous and future child outcomes. The literature that we survey estimates the economic preferences of children and links these preferences to SES, parents, peers, and field outcomes.

We view the main methodological takeaways from the experimental literature as follows. First, general patterns in economic preferences related to age or demographics have now been identified, especially for elicitation tasks that are relatively easy to explain (i.e., for dictator games, time preferences and risk preferences). Second, links between some experimental tasks and some field behaviors have been established. This is good news since it means that methods currently used are mostly appropriate for measuring the intended constructs. An outlier is competitiveness – evidence in this area is more mixed and authors have attributed these results to differences across cultures or methods. Third, results on parental transmission of preferences and impact of environment are also mixed. A commonly accepted set of methods for conducting experiments with children will be helpful in reconciling some of these diverse results.

In the introduction to this manuscript, we discussed several questions that are important for silo 2 – understanding children to understand adults. Namely, we asked about the association of preferences with family background, the role of experience in shaping preferences, and about when traits and preferences emerge and how they change over time. The literature to date has begun to

³¹ This is not the only consideration when conducting intervention studies. For example, interventions should be designed to be culturally appropriate, which can be achieved through iterative discussions with relevant stake-holders, e.g., families, teachers, school administrators in the area where the intervention will be conducted.

provide answers to these questions. For instance, as evidenced by the research summarized in Sections 5 and 6, we are learning that preferences are associated with family background but are also malleable in the sense that they evolve as children grow up. The research described in Section 8 shows that inputs from external factors such as schooling affect preferences, in the sense that (some) preferences can be affected by (some) interventions. A remaining open question here is whether such interventions affect preferences themselves, or whether they affect the behavioral strategies that are used by the children. That is, if we use an intervention to affect willingness to wait for future rewards, have we changed underlying preferences or simply provided children with tools to use to help them wait for future rewards?

Further, we have begun to answer the question about whether parents and external influences affect preferences. It is clear that SES and parental preferences are critically important since research summarized in Sections 6.1 and 6.2 shows that both SES and parental preferences are correlated with child preferences. In the case of parental preferences, the correlations are more commonly found in studies of older children and not strong in studies with younger children, suggesting that perhaps the transmission of preferences from parent to child takes time to manifest. This provides inferences into understanding why preferences of adults are widely heterogeneous. In particular, when considering inequalities by SES among adults, one important driver could be the differences in economic preferences manifested at an early age.

Finally, it is interesting to note that many preferences and behaviors are already present at a very early age. As discussed in Section 3, very young children are able to make rational decisions and make trades in markets. Further, as discussed in Section 5.4, the gender gap in competitiveness observed in adults emerges at an early age. Thus, there seems to be a component of economic preferences that arises early in life, even prior to formal schooling. This points to the importance of studying children prior to school-age, which is also discussed in List et al., (2018).

There are also a number of open questions in the related work. First, few papers focus on field outcomes that are long term. Instead, many of the existing papers discuss abstract topics or experiments in very controlled settings with short-term outcomes. A deeper understanding of how preferences affect field outcomes in the long-term – for example, how economic preferences affect human capital formation; household production; consumption; within-household behavior; and between peer behavior – would be valuable. Such work would need to be longitudinal in nature and would be valuable to understand decisions over the life course. This would lead policymakers

and researchers to place more value on economic preferences. Another open question is how preferences affect wellbeing. This does not mean that intermediate-term outcomes should be ignored; indeed, research suggests that collecting multiple intermediate outcomes can be helpful for drawing inference about long-term outcomes (e.g., Athey et al., 2019).

Second, few papers gather causal evidence. Instead, much of the work correlates SES or parenting with economic preferences. This correlational work is important, but causal inference demands either natural or field experiments to assess how investments in children affect preferences. It is also instructive to explore heterogeneous treatment effects, since how children respond to an intervention might itself be affected by preferences. Interventions designed to affect a skill or welfare (such as early childhood programs) may provide benefits in that area, but may also have unintended consequences on preferences. Understanding the consequences on preferences (as is done in Cappelen et al., 2020) is valuable, since if policies shape preferences, then they ultimately shape institutions as children become adults. Another open question is whether there are “critical periods” to intervene with children. Conducting experiments that exogenously manipulate the environment at different ages would allow researchers to answer such questions. Finally, we pointed out in this paper that children can be considered “inexperienced” subjects in many ways, hence they can be used for further study to understand how experience affects decision making, as in List (2004a).

The field is relatively new, and therefore a common methodology for conducting experiments with children has not yet been established. While Table 1 identifies what methods most papers use in each age range, there are a number of papers that diverge from the most common – and therefore potentially from the most appropriate - method. It is our hope that this paper advances our understanding of how best to conduct experiments with children and adolescents, and provides a common set of guidelines that allow experimental designs to become more consistent with one another. This will allow for more direct comparability in findings.

The field has had a historic lack of discussion with the sister sciences. Tools from psychology – such as using multiple measures and assessing test-retest reliability – can be helpful in advancing our understanding of how preference measurements should be collected. Multiple measures are especially important in cases where we expect error, and economists have thought about this issue (e.g., Heckman et al., 2013; Gillen et al., 2019). Collecting data on how other fields measure similar constructs (such as using stated-preference questionnaires for risk or time) can be

valuable for designing multiple measures. Alternatively, economists also have value to add to the sister sciences. For instance, economists use incentivized, context-free elicitation tasks that differ significantly from those used in the sister sciences. As such, several papers in the psychology literature already use economic games such as dictator games. Economists are also well trained in quantitative methods, and can impart knowledge about power, multiple hypothesis testing and related concepts to researchers in other fields.

12. References

- Abrams, Eliot, Jonathan Libgober, and John A. List, 2020. "Research Registries: Facts, Myths, and Possible Improvements," Artefactual Field Experiments 00703, The Field Experiments Website.
- Al-Ubaydli, Omar, John A. List, and Dana L. Suskind. 2017b. "What Can We Learn from Experiments? Understanding the Threats to the Scalability of Experimental Results." *American Economic Review* 107 (5): 282–86.
- Al-Ubaydli, Omar, John A. List, Danielle Lore, and Dana Suskind. 2017a. "Scaling for Economists: Lessons from the Non-Adherence Problem in the Medical Literature." *Journal of Economic Perspectives* 31 (4): 125–44.
- Al-Ubaydli, Omar, John List, and Dana Suskind. 2020. "The Science of Using Science: Towards an Understanding of the Threats to Scaling Experiments." *International Economic Review*, 61(4): 1387-1409.
- Al-Ubaydli, Omar, Min Sok Lee, John A. List, Claire Mackevicius, and Dana Suskind. 2021. "How Can Experiments Play a Greater Role in Public Policy? 12 Proposals from an Economic Model of Scaling." *Behavioural Public Policy*, 5(1): 2-49.
- Alan, Sule, and Seda Ertac. 2018a. "Fostering Patience in the Classroom: Results from Randomized Educational Intervention." *Journal of Political Economy* 126 (5): 1865–1911.
- Alan, Sule, and Seda Ertac. 2018b. "Mitigating the Gender Gap in the Willingness to Compete: Evidence from a Randomized Field Experiment." *Journal of the European Economic Association* 17 (4): 1147–85.
- Alan, Sule, Nazli Baydar, Teodora Boneva, Thomas F. Crossley, and Seda Ertac. 2017. "Transmission of Risk Preferences from Mothers to Daughters." *Journal of Economic Behavior & Organization* 134: 60–77.
- Alan, Sule, Teodora Boneva, and Seda Ertac. 2019. "Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit." *The Quarterly Journal of Economics* 134 (3).
- Almås, Ingvild, Alexander W. Cappelen, and Bertil Tungodden. 2019. "Cutthroat Capitalism versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians?" *Journal of Political Economy*.
- Almås, Ingvild, Alexander W. Cappelen, Kjell G. Salvanes, Erik Ø. Sørensen, and Bertil Tungodden. 2016. "Fairness and Family Background." *Politics, Philosophy & Economics* 16 (2): 117–31.
- Almond, Douglas, and Janet Currie. 2011. "Killing Me Softly: The Fetal Origins Hypothesis." *Journal of Economic Perspectives* 25 (3): 153–72.

- Almond, Douglas, Janet Currie, and Valentina Duque. 2018. "Childhood Circumstances and Adult Outcomes: Act II." *Journal of Economic Literature* 56 (4): 1360–1446.
- Andersen, Steffen, Seda Ertac, Uri Gneezy, John A. List, and Sandra Maximiano. 2013. "Gender, Competitiveness, and Socialization at a Young Age: Evidence From a Matrilineal and a Patriarchal Society." *Review of Economics and Statistics* 95 (4): 1438–43.
- Anderson, Lisa R., and Jennifer M. Mellor. 2008. "Predicting Health Behaviors with an Experimental Measure of Risk Preference." *Journal of Health Economics* 27 (5): 1260–74.
- Andreoni, James, Amalia Di Girolamo, John List, Claire Mackevicius, and Anya Samek. 2019a. "Risk Preferences of Children and Adolescents in Relation to Gender, Cognitive Skills, Soft Skills, and Executive Functions." *National Bureau of Economic Research*, no. w25723.
- Andreoni, James, Anya Samek, John A List, and Charles Sprenger. 2017. "Field Experiments on the Development of Time Preferences." *The Field Experiments Website*, no. 00615.
- Andreoni, James, Michael Kuhn, John List, Anya Samek, Kevin Sokal, and Charles Sprenger. 2019b. "Toward an Understanding of the Development of Time Preferences: Evidence from Field Experiments." *National Bureau of Economic Research*, no. w25590.
- Angerer, Silvia, Daniela Glätzle-Rützler, Philipp Lergetporer, and Matthias Sutter. 2015a. "Donations, Risk Attitudes and Time Preferences: A Study on Altruism in Primary School Children." *Journal of Economic Behavior & Organization* 115: 67–74.
- Angerer, Silvia, Philipp Lergetporer, Daniela Glätzle-Rützler, and Matthias Sutter. 2015b. "How to Measure Time Preferences in Children: a Comparison of Two Methods." *Journal of the Economic Science Association* 1 (2): 158–69.
- Angerer Silvia, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter. 2016. "Cooperation and discrimination within and across language borders: Evidence from children in a bilingual city" *European Economic Review* 90: 254-264
- Antrobus, Emma, Deborah A Cobb-Clark, Lorraine Mazerolle, and Agnieszka Tymula. 2017. "The Risk and Time Preferences of Young Truants and Their Parents."
- Athey, Susan, Raj Chetty, Guido W. Imbens, and Hyunseung Kang. 2019 "The surrogate index: Combining short-term proxies to estimate long-term treatment effects more rapidly and precisely." *National Bureau of Economic Research* no. w26463.
- Arrow, Kenneth J. 1972. "Gifts and Exchanges." *Philosophy & Public Affairs*, 343–62.
- Attanasio, Orazio P., Costas Meghir, and Ana Santiago. 2012. "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate Progreso." *The Review of Economic Studies* 79 (1): 37-66.
- Attanasio, O., S. Cattan, E. Fitzsimons, C. Meghir, and M. Rubio-Codina. 2020. "Estimating the production function for human capital: results from a randomized controlled trial in Colombia." *American Economic Review* 110 (1), 48–85.
- Badarudeen, Sameer, and Sanjeev Sabharwal. 2010. "Assessing Readability of Patient Education Materials: Current Role in Orthopaedics." *Clinical Orthopaedics and Related Research* 468 (10): 2572–80.
- Barash, Jori, Isabelle Brocas, Juan D. Carrillo, and Niree Kodaverdian. 2019. "Heuristic to Bayesian: The Evolution of Reasoning from Childhood to Adulthood." *Journal of Economic Behavior & Organization* 159: 305–22.
- Bauer, Michal, Alessandra Cassar, Julie Chytilová, and Joseph Henrich. 2014a. "War's Enduring Effects on the Development of Egalitarian Motivations and In-Group Biases." *Psychological Science* 25 (1): 47–57.

- Bauer, Michal, Julie Chytilová, and Barbara Pertold-Gebicka. 2014b. "Parental Background and Other-Regarding Preferences in Children." *Experimental Economics* 17 (1): 24–46.
- Bekkers, Rene. 2017. "Do Two Eyes See More Generosity Than One?"
- Ben-Ner, Avner, John A. List, Louis Putterman, and Anya S. Samek. 2017. "Learned Generosity? An Artefactual Field Experiment with Parents and Their Children." *Journal of Economic Behavior & Organization* 143: 28–44.
- Benenson, Joyce F., Joanna Pascoe, and Nicola Radmore. 2007. "Childrens Altruistic Behavior in the Dictator Game." *Evolution and Human Behavior* 28 (3): 168–75.
- Benjamin, Daniel J., David Cesarini, Christopher F. Chabris, Edward L. Glaeser, David I. Laibson, Gene Age, Vilmondur Guðnason et al. "The promises and pitfalls of genoconomics." (2011).
- Benjamin, Daniel J., David Laibson, Walter Mischel, Philip K. Peake, Yuichi Shoda, Alexandra Steiny Wellsjo, and Nicole L. Wilson. 2019. "Predicting Mid-Life Capital Formation with Pre-School Delay of Gratification and Life-Course Measures of Self-Regulation." *Journal of Economic Behavior & Organization*.
- Benjamin, Daniel J., Sebastian A. Brown, and Jesse M. Shapiro. 2013. "Who Is Behavioral? Cognitive Ability and Anomalous Preferences." *Journal of the European Economic Association* 11 (6): 1231–55.
- Bereby-Meyer, Yoella, and Shelly Fiks. 2013. "Changes in Negative Reciprocity as a Function of Age." *Journal of Behavioral Decision Making* 26 (4): 397–403.
- Bergen, Doris. 2002. "The Role of Pretend Play in Children's Cognitive Development." *Early Childhood Research & Practice* 4 (1): n1.
- Bertrand, Marianne, and Kevin F. Hallock. 2001. "The Gender Gap in Top Corporate Jobs." *ILR Review* 55 (1): 3–21.
- Bertrand, Marianne. 2009. "CEOs." *Annual Review of Economics* 1 (1): 121–50.
- Bettinger, Eric, and Robert Slonim. 2006. "Using Experimental Economics to Measure the Effects of a Natural Educational Experiment on Altruism." *Journal of Public Economics* 90 (8-9): 1625–48.
- Bettinger, Eric, and Robert Slonim. 2007. "Patience among Children." *Journal of Public Economics* 91 (1-2): 343–63.
- Bhuller, Manudeep, Magne Mogstad, and Kjell G. Salvanes. "Life-cycle earnings, education premiums, and internal rates of return." *Journal of Labor Economics* 35, no. 4 (2017): 993-1030.
- Bigelow, James, and Amy Poremba. 2014. "Achilles' Ear? Inferior Human Short-Term and Recognition Memory in the Auditory Modality." *PLoS ONE* 9 (2): e89914.
- Bindra, Parampreet Christopher, Daniela Rützler, and Philipp Lergetporer. 2018. "Discrimination at Young Age: Experimental Evidence from Preschool Children."
- Bisin, Alberto, and Thierry Verdier. 2001. "The Economics of Cultural Transmission and the Dynamics of Preferences." *Journal of Economic Theory* 97 (2): 298–319.
- Blake, Peter R., and Katherine McAuliffe. 2011. "'I Had So Much It Didn't Seem Fair': Eight-Year-Olds Reject Two Forms of Inequity." *Cognition* 120 (2): 215-224.
- Blau, Francine D., Marianne A. Ferber, and Anne E. Winkler. 2010. *The Economics of Women, Men, and Work*. 6th ed. Upper Saddle River, NJ: Prentice Hall.
- Boneva, Teodora, and Christopher Rauh. "Parental beliefs about returns to educational investments—the later the better?." *Journal of the European Economic Association* 16, no. 6 (2018): 1669-1711.

- Booth, Alison, and Patrick Nolen. 2012. "Choosing to Compete: How Different Are Girls and Boys?" *Journal of Economic Behavior & Organization* 81 (2): 542–55.
- Boswell, Nikki, Rebecca Byrne, and Peter S.W. Davies. 2019. "An Examination of Children's Eating Behaviours as Mediators of the Relationship Between Parents' Feeding Practices and Early Childhood Body Mass Index z-Scores." *Obesity science & practice* 5 (2): 168–176.
- Bradford, W. David, Charles Courtemanche, Garth Heutel, Patrick McAlvanah, and Christopher Ruhm. 2017. "Time Preferences and Consumer Behavior." *Journal of Risk and Uncertainty* 55 (2-3): 119–45.
- Bradford, W. David. 2010. "The Association Between Individual Time Preferences and Health Maintenance Habits." *Medical Decision Making* 30 (1): 99–112.
- Brocas, Isabelle, and Juan D. Carrillo. 2018. "The Determinants of Strategic Thinking in Preschool Children." *PLoS One* 13 (5).
- Brocas, Isabelle, and Juan D. Carrillo. 2019a. "Money in the schoolyard: young children use commodities as an indirect medium of exchange." Working paper.
- Brocas, Isabelle, and Juan D. Carrillo. 2019b. "Steps of reasoning in children and adolescents." *Working Paper*.
- Brocas, Isabelle, Juan D. Carillo, and Niree Kodaverdian. 2017. "Altruism and Strategic Giving in Children and Adolescents."
- Brocas, Isabelle, Juan D. Carillo, T. Dalton Combs, and Niree Kodaverdian. 2016. "Value-Based Decision-Making: A New Developmental Paradigm."
- Brocas, Isabelle, Juan D. Carrillo, T. Dalton Combs, and Niree Kodaverdian. 2019. "The Development of Consistent Decision-Making across Economic Domains." *Games and Economic Behavior* 116: 217–40.
- Brock, William A., and Steven N. Durlauf. 2001. "Discrete Choice with Social Interactions." *The Review of Economic Studies* 68 (2): 235–60.
- Brock, William A., and Steven N. Durlauf. 2002. "A Multinomial-Choice Model of Neighborhood Effects." *American Economic Review* 92 (2): 298–303.
- Broome, Marion E., and Marti Rice. 2004. "Incentives for Children in Research." *Journal of Nursing Scholarship* 36 (2): 167–72.
- Brosig-Koch, Jeannette, Timo Heinrich, and Christoph Helbach. 2015. "Exploring the Capability to Reason Backwards: An Experimental Study with Children, Adolescents, and Young Adults." *European Economic Review* 74: 286–302.
- Bunch, Katie M., Glenda Andrews, and Graeme S. Halford. 2007. "Complexity Effects on the Childrens Gambling Task." *Cognitive Development* 22 (3): 376–83.
- Burns, Zach, Andrew Chiu, and George Wu. 2010. "Overweighting of Small Probabilities." *Wiley Encyclopedia of Operations Research and Management Science*.
- Busby, Janie, and Thomas Suddendorf. 2005. "Recalling Yesterday and Predicting Tomorrow." *Cognitive Development* 20 (3): 362–72.
- Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. 2014. "Gender, Competitiveness, and Career Choices." *The Quarterly Journal of Economics* 129 (3): 1409–47.
- Cadena, Brian C., and Benjamin J. Keys. 2015. "Human Capital and the Lifetime Costs of Impatience." *American Economic Journal: Economic Policy* 7 (3): 126–53.
- California Department of Education. 2013. "California Common Core State Standards: Mathematics."

- Callen, Michael, Mohammad Isaqzadeh, James D. Long, and Charles D. Sprenger. 2014. "Violence and Risk Preference: Experimental Evidence from Afghanistan." *American Economic Review* 104 (1): 123–48.
- Campbell, Jamie ID. 2005. "Handbook of Mathematical Cognition." *Psychology Press*.
- Cappelen, Alexander W., John A. List, Anya S. Samek, and Bertil Tungodden. Forthcoming. "The Effect of Early Childhood Education on Social Preferences." *Journal of Political Economy*, 2020.
- Carper, Doris. 1942. "Seeing Numbers as Groups in Primary-Grade Arithmetic." *The Elementary School Journal* 43 (3): 166–70.
- Carroll, John. B. 1962. "The prediction of success in intensive foreign language training." In R. Glaser (Ed.), *Training research and education* (pp. 87–136). University of Pitt. Press.
- Carrillo, Bladimir. "Present Bias and Underinvestment in Education? Long-run Effects of Childhood Exposure to Booms in Colombia." *Journal of Labor Economics* 38, no. 4 (2020): 000-000.
- Castillo, Marco, Jeffrey L. Jordan, and Ragan Petrie. 2018. "Children's Rationality, Risk Attitudes and Field Behavior." *European Economic Review* 102: 62–81.
- Castillo, Marco, Jeffrey L. Jordan, and Ragan Petrie. 2019a. "Discount Rates of Children and High School Graduation." *The Economic Journal* 129 (619): 1153–81. doi:10.1111/eoj.12574.
- Castillo, Marco, John A List, Ragan Petrie, and Anya Samek. 2020. "Detecting Drivers of Behavior at an Early Age: Evidence from a Longitudinal Field Experiment." *NBER Working Paper* 28288.
- Castillo, Marco, Paul J. Ferraro, Jeffrey L. Jordan, and Ragan Petrie. 2011. "The Today and Tomorrow of Kids: Time Preferences and Educational Outcomes of Children." *Journal of Public Economics* 95 (11-12): 1377–85.
- Castillo, Marco, Ragan Petrie, and Maximo Torero. 2019b. "Causal Evidence That Perinatal Conditions Affect the Risk Preferences of Adults." *Working Paper*.
- Castillo, Marco, 2020, "Negative Childhood Experiences and Risk Aversion: Evidence from Children Exposed to Domestic Violence," *IZA DP No. 13320*
- Chabris, Christopher F., David Laibson, Carrie L. Morris, Jonathon P. Schuldt, and Dmitry Taubinsky. 2008. "Individual Laboratory-Measured Discount Rates Predict Field Behavior." *Journal of Risk and Uncertainty* 37 (2-3): 237.
- Chamberlin, Edward H. 1948. "An Experimental Imperfect Market." *Journal of Political Economy* 56 (2): 95–108.
- Chapman, Gretchen B., and Elliot J. Coups. 1999. "Time Preferences and Preventive Health Behavior." *Medical Decision Making* 19 (3): 307–14.
- Charness, Gary, John A. List, Aldo Rustichini, Anya S. Samek, and Jeroen Van De Ven. 2019. "Theory of Mind among Disadvantaged Children: Evidence from a Field Experiment." *Journal of Economic Behavior & Organization* 166: 174–94.
- Chen, Jingnan, Daniel Houser, Natalia Montinari, and Marco Piovesan. 2016. "Beware of Popular Kids Bearing Gifts: A Framed Field Experiment." *Journal of Economic Behavior & Organization* 132: 104–20.
- Chernyak, Nadia, Beth Sandham, Paul L. Harris, and Sara Cordes. 2016. "Numerical Cognition Explains Age-Related Changes in Third-Party Fairness." *Developmental Psychology* 52 (10): 1555–62.

- Chowdhury, Shyamal, Matthias Sutter, and Klaus Zimmermann. 2018. "Evaluating Intergenerational Persistence of Economic Preferences: A Large Scale Experiment with Families in Bangladesh." *SSRN Electronic Journal*.
- Chyn, Eric. 2018. "Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children." *American Economic Review* 108 (10): 3028–56.
- Clements, Douglas H., Julie Sarama, and Beth L. Macdonald. 2019. "Subitizing: The Neglected Quantifier." *Constructing Number Research in Mathematics Education*, 13–45.
- Cohen, Michael A., Todd S. Horowitz, and Jeremy M. Wolfe. 2009. "Auditory Recognition Memory Is Inferior to Visual Recognition Memory." *Proceedings of the National Academy of Sciences* 106 (14): 6008–10.
- Coleman, James S. 1988. "Social capital in the creation of human capital." *American Journal of Sociology* 94, S95-S120.
- Cotton, Christopher, Brent R. Hickman, John A. List, Joseph Price, and Sutanuka Roy. 2020. "Productivity Versus Motivation in Adolescent Human Capital Production: Evidence from a Structurally-Motivated Field Experiment." Framed Field Experiments 00719 working paper, The Field Experiments Website and *University of Chicago, Becker Friedman Institute for Economics Working Paper* 2020-150.
- Cowell, Jason M., Anya S. Samek, John List, and Jean Decety. 2015. "The Curious Relation between Theory of Mind and Sharing in Preschool Age Children." *PLoS One* 10 (2): e0117947.
- Cox, James C., John A. List, Michael Price, Vjollca Sadiraj, and Anya S. Samek. 2016. "Moral Costs and Rational Choice: Theory and Experimental Evidence." *National Bureau of Economic Research*, no. w22234.
- Cunha, Flavio, and James Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97 (2): 31–47.
- Currie, Janet. 2001. "Early Childhood Education Programs." *Journal of Economic Perspectives* 15 (2): 213–38.
- Czermak, Simon, Francesco Feri, Daniela Glätzle-Rützler, and Matthias Sutter. 2016. "How Strategic Are Children and Adolescents? Experimental Evidence from Normal-Form Games." *Journal of Economic Behavior & Organization* 128: 265–85.
- Czibor, Eszter, David Jimenez-Gomez, and John A. List. 2019. "The Dozen Things Experimental Economists Should Do (More of)." *National Bureau of Economic Research*, no. w25451.
- Dajani, Dina R., and Lucina Q. Uddin. 2015. "Demystifying Cognitive Flexibility: Implications for Clinical and Developmental Neuroscience." *Trends in Neurosciences* 38 (9): 571–78.
- Datar, Ashlesha, Nancy Nicosia, and Anya Samek. 2019. "The Moderating Effects of Time and Risk Preferences on the Relationship Between Neighborhood Environments and Obesity: Evidence from a Natural Experiment. *Work in Progress*.
- Davis, Douglas D., and Charles A. Holt. 1993. *Experimental Economics*. Princeton University Press.
- Deckers, Thomas, Armin Falk, Fabian Kosse, and Hannah Schildberg-Hörisch. 2015. "How Does Socio-Economic Status Shape a Child's Personality?"
- Del Boca, Daniela, Christopher Flinn, Ewout Verriest, and Matthew Wiswall. 2019. "Actors in the Child Development Process." *National Bureau of Economic Research*, no. w25596.
- DellaVigna, Stefano, and M. Daniele Paserman. 2005. "Job Search and Impatience." *Journal of Labor Economics* 23 (3): 527–88.

- DellaVigna, Stefano, John A. List, Ulrike Malmendier, and Gautam Rao. 2013. "The Importance of Being Marginal: Gender Differences in Generosity." *American Economic Review*, 103 (3): 586-90.
- Devine, R., and C. Hughes. 2018. "Family Correlates of False Belief Understanding in Early Childhood: A Meta-Analysis." *Child Development*, 89: 971-987.
- DeVon, Holli A., Michelle E. Block, Patricia Moyle-Wright, Diane M. Ernst, Susan J. Hayden, Deborah J. Lazzara, Suzanne M. Savoy, and Elizabeth Kostas-Polston. 2007. "A Psychometric Toolbox for Testing Validity and Reliability." *Journal of Nursing Scholarship* 39 (2): 155-64.
- Dixon, Wallace E., Brenda J. Salley, and Andrea D. Clements. 2006. "Temperament, Distraction, and Learning in Toddlerhood." *Infant Behavior and Development* 29 (3): 342-57.
- Dizon-Ross, Rebecca. 2019. "Parents' beliefs about their children's academic ability: Implications for educational investments." *American Economic Review* 109, no. 8 : 2728-65.
- Doepke, Matthias, and Fabrizio Zilibotti. 2017. "Parenting With Style: Altruism and Paternalism in Intergenerational Preference Transmission." *Econometrica* 85 (5): 1331-71.
- Dohmen, T., A. Falk, D. Huffman, and U. Sunde. 2011. "The Intergenerational Transmission of Risk and Trust Attitudes." *The Review of Economic Studies* 79 (2): 645-77.
- Dreber, Anna, Emma Von Essen, and Eva Ranehill. 2011. "Outrunning the Gender Gap—Boys and Girls Compete Equally." *Experimental Economics* 14 (4): 567-82.
- Dreber, Anna, Emma Von Essen, and Eva Ranehill. 2014. "Gender and Competition in Adolescence: Task Matters." *Experimental Economics* 17 (1): 154-72.
- Dumontheil, Iroise, Ian A. Apperly, and Sarah-Jayne Blakemore. "Online usage of theory of mind continues to develop in late adolescence." *Developmental science* 13, no. 2 (2010): 331-338.
- Duncan, Greg J, and Katherine Magnuson. 2013. "Investing in Preschool Programs." *Journal of Economic Perspectives* 27 (2): 109-32.
- Durlauf, Steven N. 2014. "Neighborhood Effects." *Handbook of Regional and Urban Economics* 4: 2173-2242.
- Dweck, Carol. 2015. "Carol Dweck revisits the growth mindset." *Education Week* 35, no. 5 : 20-24.
- Dynarski, Susan, C.J. Libassi, Katherine Micheltmore, and Stephanie Owen. 2018. "Closing the Gap: The Effect of a Targeted, Tuition-Free Promise on College Choices of High-Achieving, Low-Income Students." *National Bureau of Economic Research*, no. w25349.
- Eckel, Catherine C., Philip J. Grossman, Cathleen A. Johnson, Angela C. M. De Oliveira, Christian Rojas, and Rick K. Wilson. 2012. "School Environment and Risk Preferences: Experimental Evidence." *Journal of Risk and Uncertainty* 45 (3): 265-92.
- Evans, Anthony M., Ursula Athenstaedt, and Joachim I. Krueger. 2013. "The Development of Trust and Altruism during Childhood." *Journal of Economic Psychology* 36: 82-95.
- Falk, Armin, and Fabian Kosse. 2016. "Early Childhood Environment, Breastfeeding and the Formation of Preferences." *SSRN Electronic Journal*.
- Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde. 2018. "Global Evidence on Economic Preferences." *The Quarterly Journal of Economics* 133 (4): 1645-92.
- Fe, Eduardo, David Gill, and Victoria Prowse. *Cognitive skills, strategic sophistication, and life outcomes*. Working Paper, 2020.

- Fehr, Ernst, Daniela Glätzle-Rützler, and Matthias Sutter. 2013. "The Development of Egalitarianism, Altruism, Spite and Parochialism in Childhood and Adolescence." *European Economic Review* 64: 369–83.
- Fehr, Ernst, Helen Bernhard, and Bettina Rockenbach. 2008. "Egalitarianism in Young Children." *Nature* 454 (7208): 1079.
- Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics* 10 (2): 171–78.
- Fischbein, H. 1975. "The Intuitive Sources of Probabilistic Thinking in Children." *Springer Science & Business Media* 85.
- Fisher, Anna V., Karrie E. Godwin, and Howard Seltman. 2014. "Visual Environment, Attention Allocation, and Learning in Young Children." *Psychological Science* 25 (7): 1362–70.
- Fitzgerald, Jill, and Timothy Shanahan. "Reading and writing relations and their development." *Educational Psychologist* 35, no. 1 (2000): 39-50.
- Flory, Jeffrey A., Andreas Leibbrandt, and John A. List. 2015. "Do Competitive Workplaces Deter Female Workers? A Large-Scale Natural Field Experiment on Job Entry Decisions." *The Review of Economic Studies* 82 (1): 122–155.
- Flory, Jeffrey A. & Gneezy, Uri & Leonard, Kenneth L. & List, John A., 2018. "Gender, age, and competition: A disappearing gap?," Journal of Economic Behavior & Organization, 150(C): 256-276.
- Foss, Elizabeth, Allison Druin, Robin Brewer, Phillip Lo, Luis Sanchez, Evan Golub, and Hilary Hutchinson. 2011. "Childrens Search Roles at Home: Implications for Designers, Researchers, Educators, and Parents." *Journal of the American Society for Information Science and Technology* 63 (3): 558–73.
- Fréchette, Guillaume R. 2016. "Experimental Economics across Subject Populations." *The Handbook of Experimental Economics* 2: 435–80.
- Fryer Jr, Roland G. 2017. "The Production of Human Capital in Developed Countries." *Handbook of Economic Field Experiments* 2: 95–322.
- Fryer Jr, Roland G., Steven D. Levitt, and John A. List. 2015. "Parental Incentives and Early Childhood Achievement: A Field Experiment in Chicago Heights." *National Bureau of Economic Research*, no. w21477.
- Fryer Jr, Roland G., Steven D. Levitt, John A. List, and Anya S. Samek. 2020. "Towards an Understanding of What Works in Preschool Education." *Work in Progress*.
- Gaspelin, Nicholas, Tessa Margett-Jordan, and Eric Ruthruff. 2015. "Susceptible to Distraction: Children Lack Top-down Control over Spatial Attention Capture." *Psychonomic Bulletin & Review* 22 (2): 461–68.
- Gathercole, Susan E., Susan J. Pickering, Benjamin Ambridge, and Hannah Wearing. 2004. "The Structure of Working Memory From 4 to 15 Years of Age." *Developmental Psychology* 40 (2): 177.
- Gaviria, Alejandro, and Steven Raphael. 2001. "School-Based Peer Effects and Juvenile Behavior." *Review of Economics and Statistics* 83 (2): 257–68.
- Gelman, Rochel, and Charles R. Gallistel. 1978. "The Child's Understanding of Number." *Harvard University Press*.
- Gershon, Richard C., Molly V. Wagster, Hugh C. Hendrie, Nathan A. Fox, Karon F. Cook, and Cindy J. Nowinski. 2013. "NIH Toolbox for Assessment of Neurological and Behavioral Function." *Neurology* 80 (11 Supplement 3): S2–S6.

- Gillen, Ben, Erik Snowberg, and Leeat Yariv. 2019. "Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study." *Journal of Political Economy* 127 (4): 1826–63.
- Gneezy, Uri, and Aldo Rustichini. 2004. "Gender and Competition at a Young Age." *American Economic Review* 94 (2): 377–81.
- Gneezy, Uri, Kenneth L. Leonard, and John A. List. 2009. "Gender Differences in Competition: Evidence From a Matrilineal and a Patriarchal Society." *Econometrica* 77 (5): 1637–64.
- Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. 2003. "Performance in Competitive Environments: Gender Differences." *The Quarterly Journal of Economics* 118 (3): 1049–74.
- Golsteyn, Bart H.H., Hans Grönqvist, and Lena Lindahl. 2014. "Adolescent Time Preferences Predict Lifetime Outcomes." *The Economic Journal* 124 (580).
- Halevy, Yoram. 2007. "Ellsberg Revisited: An Experimental Study." *Econometrica* 75 (2): 503–36.
- Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein. 2014. "Learning Through Noticing: Theory and Evidence from a Field Experiment." *The Quarterly Journal of Economics* 129 (3): 1311–53.
- Harbaugh, William T., Kate Krause, and Lise Vesterlund. 2002. "Risk Attitudes of Children and Adults: Choices over Small and Large Probability Gains and Losses." *Experimental Economics* 5 (1): 53–84.
- Harbaugh, William T., Kate Krause, and Timothy R. Berry. 2001. "GARP for Kids: On the Development of Rational Choice Behavior." *American Economic Review* 91 (5): 1539–45.
- Harbaugh, William T., Kate Krause, Steven G. Liday, and Lise Vesterlund. 2003. "Trust in Children." *Trust and Reciprocity: Interdisciplinary Lessons from Experimental Research*, 302–22.
- Harrison, Glenn W., and John A. List. 2004. "Field Experiments." *Journal of Economic Literature* 42 (4): 1009–55.
- Harrison, Glenn W., Morten I. Lau, and E. Elisabet Rutström. 2007. "Estimating Risk Attitudes in Denmark: A Field Experiment." *Scandinavian Journal of Economics* 109 (2): 341–68.
- Heckman, James J. 2006. "Skill Formation and the Economics of Investing in Disadvantaged Children." *Science* 312 (5782): 1900–1902.
- Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz. 2010. "The Rate of Return to the HighScope Perry Preschool Program." *Journal of Public Economics* 94 (1-2): 114–28.
- Heckman, James, and Ganesh Karapakula. 2019. "Intergenerational and Intragenerational Externalities of the Perry Preschool Project." *National Bureau of Economic Research*, no. w25889.
- Heckman, James, Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review* 103 (6): 2052–86.
- Hirsch, Jerry. 1963. "Behavior Genetics and Individuality Understood." *Science* 142 (3598): 1436–42.
- Houser, Daniel, John List, Marco Piovesan, Anya Savikhin Samek, and Joachim Winter. 2015. "On the Origins of Dishonesty: From Parents to Children." *National Bureau of Economic Research*, no. W20897.

- Houser, Daniel, Natalia Montinari, and Marco Piovesan. 2012. "Private and Public Decisions in Social Dilemmas: Evidence from Children's Behavior." *PLoS ONE* 7 (8): e41568.
- Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. "Long-run impacts of childhood access to the safety net." *American Economic Review* 106 (4): 903-34.
- Ifcher, John, and Homa Zarghamee. 2016. "Do Gender-Variant Preferences For Competition Persist In The Absence Of Performance?" *Economic Inquiry* 54 (4): 1918-30.
- IJzendoorn, Marinus H. Van, Marian J. Bakermans-Kranenburg, and Richard P. Ebstein. 2011. "Methylation Matters in Child Development: Toward Developmental Behavioral Epigenetics." *Child Development Perspectives* 5 (4): 305-10.
- Ikeda, Shinsuke, Myong-Il Kang, and Fumio Ohtake. 2010. "Hyperbolic Discounting, the Sign Effect, and the Body Mass Index." *Journal of Health Economics* 29 (2): 268-84.
- Jacob, Brian A. 2004. "Public Housing, Housing Vouchers, and Student Achievement: Evidence from Public Housing Demolitions in Chicago." *American Economic Review* 94 (1): 233-58.
- Joensen, Juanna Shroter, John List, Anya Samek, Haruka Uchida. 2021. "Using a Field Experiment to Understand Skill Formation in Adolescent Years" *Work in progress*
- Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (2): 263-91.
- Kautz, T, Heckman, JJ. et al. 2014, "Fostering and Measuring Skills" Paris, OECD
- Khachatryan, Karen, Anna Dreber, Emma Von Essen, and Eva Ranehill. 2015. "Gender and preferences at a young age: Evidence from Armenia." *Journal of Economic Behavior & Organization* 118: 318-332.
- Khadjavi, Menusch, and Andreas Nicklisch. 2018. "Parents' ambitions and children's competitiveness." *Journal of Economic Psychology* 67: 87-102.
- Kidd, Celeste, Holly Palmeri, and Richard N. Aslin. 2013. "Rational Snacking: Young Children's Decision-Making on the Marshmallow Task Is Moderated by Beliefs about Environmental Reliability." *Cognition* 126 (1): 109-14.
- Kim, Hojin I., and Scott P. Johnson. 2013. "Do Young Infants Prefer an Infant-Directed Face or a Happy Face?" *International Journal of Behavioral Development* 37 (2): 125-30.
- Kimball, Miles S., Claudia R. Sahm, and Matthew D. Shapiro. 2009. "Risk preferences in the PSID: individual imputations and family covariation." *American Economic Review* 99 (2): 363-68.
- Kosse, Fabian, and Friedhelm Pfeiffer. 2012. "Impatience among preschool children and their mothers." *Economics Letters* 115 (3): 493-495.
- Kosse, Fabian, Deckers, Thomas, Pinger, Pia, Schildberg-Hoerisch, Hannah, and Falk, Armin. 2019. "The Formation of Prosociality: Causal Evidence on the Role of Social Environment." *Journal of Political Economy*.
- Kremer, Michael. 2003. "Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons." *American Economic Review* 93 (2): 102-6.
- Lavecchia, Adam, Heidi Liu, and Philip Oreopoulos. 2016. "Chapter 1 - Behavioral Economics of Education: Progress and Possibilities." *Handbook of the Economics of Education* 5: 1-74.
- Ledyard, John O., John H. Kagel, and Alvin E. Roth. "Handbook of experimental economics." *Public Goods: A Survey of Experimental Research* (1995): 111-194.
- Lee, James J., Robbee Wedow, Aysu Okbay, Edward Kong, Omeed Maghazian, Meghan Zacher, Tuan Anh Nguyen-Viet et al. "Gene discovery and polygenic prediction from a genome-

- wide association study of educational attainment in 1.1 million individuals." *Nature genetics* 50, no. 8 (2018): 1112-1121.
- Leimgruber, Kristin L., Alex Shaw, Laurie R. Santos, and Kristina R. Olson. 2012. "Young Children Are More Generous When Others Are Aware of Their Actions." *PLoS ONE* 7 (10): e48292.
- Leonard, Tammy, Kerem Shuval, Angela De Oliveira, Celette Sugg Skinner, Catherine Eckel, and James C. Murdoch. 2013. "Health Behavior and Behavioral Economics: Economic Preferences and Physical Activity Stages of Change in a Low-Income African-American Community." *American Journal of Health Promotion* 27 (4): 211–21.
- Levitt, Steven D., John A. List, and Sally Sadoff. 2016a. "The Effect of Performance-Based Incentives on Educational Achievement: Evidence from a Randomized Experiment," NBER Working Paper No. 22107.
- Levitt, Steven D., John A. List, Susanne Neckermann, and Sally Sadoff. 2016b. "The behavioralist goes to school: Leveraging behavioral economics to improve educational performance." *American Economic Journal: Economic Policy* 8, no. 4: 183-219.
- Lerner, Eva, Bernhard Streicher, Rainer Sachs, Martina Raue, and Dieter Frey. 2016. "Thinking Concretely Increases the Perceived Likelihood of Risks: The Effect of Construal Level on Risk Estimation." *Risk Analysis* 36 (3): 623–37.
- Leung, Christy Y.Y., Marc W. Hernandez, Dana L. Suskind. 2020. Enriching home language environment among families from low-SES backgrounds: A randomized controlled trial of a home visiting curriculum, *Early Childhood Research Quarterly*, 50(1): 24-35.
- Liaw, Fong-Ruey, and Jeanne Brooks-Gunn. 1994. "Cumulative Familial Risks and Low-Birthweight Childrens Cognitive and Behavioral Development." *Journal of Clinical Child Psychology* 23 (4): 360–272.
- Lickliter, Robert, and David C. Witherington. 2017. "Towards a Truly Developmental Epigenetics." *Human Development* 60 (2-3): 124–38.
- Linnér, Richard Karlsson, Pietro Biroli, Edward Kong, S. Fleur W. Meddens, Robbee Wedow, Mark Alan Fontana, Maël Lebreton et al. "Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences." *Nature genetics* 51, no. 2 (2019): 245-257.
- List, John A. 2002a. "Preference Reversals of a Different Kind," *American Economic Review*, 92(5), 1636-1643.
- List, John A. 2002b. "Testing Neoclassical Competitive Market Theory in the Field." *Proceedings of the National Academy of Sciences* 99 (24): 15827–30.
- List, John A. 2003. "Does Market Experience Eliminate Market Anomalies?" *The Quarterly Journal of Economics* 118 (1): 41–71.
- List, John A. A. 2004a. "Testing Neoclassical Competitive Theory in Multilateral Decentralized Markets." *Journal of Political Economy* 112 (5): 1131–56.
- List, John A. 2004b. "Young, Selfish, and Male: Field Evidence of Social Preferences." *Economic Journal* 114: 121-149.
- List, John A. 2011. "Why Economists Should Conduct Field Experiments and 14 Tips for Pulling One Off." *Journal of Economic Perspectives* 25 (3): 3–16.
- List, John A., and Anya C. Samak. 2013. "Exploring the origins of charitable acts: Evidence from an artefactual field experiment with young children." *Economics Letters* 118, (3): 431-434.

- List, John A., and Anya S. Samek. 2015. "The Behavioralist as Nutritionist: Leveraging Behavioral Economics to Improve Child Food Choice and Consumption." *Journal of Health Economics* 39: 135–46.
- List, John A., and Daniel L. Millimet. 2008. "The Market: Catalyst for Rationality and Filter of Irrationality." *The B.E. Journal of Economic Analysis & Policy* 8 (1).
- List, John A., and Jason F. Shogren. 1998. "Calibration of the Difference between Actual and Hypothetical Valuations in a Field Experiment." *Journal of Economic Behavior & Organization* 37 (2): 193–205.
- List, John A., Anya S. Samek, and Dana L. Suskind. 2018. "Combining Behavioral Economics and Field Experiments to Reimagine Early Childhood Education." *Behavioural Public Policy* 2 (1): 1–21.
- List, John A., Fatemeh Momeni, and Yves Zenou. 2020. "The Social Side of Early Human Capital Formation: Using a Field Experiment to Estimate the Causal Impact of Neighbors." Natural Field Experiments 00722 working paper, The Field Experiments Website.
- List, Annika J., John A. List, and Anya Samek. 2017. "Discrimination among pre-school children: Field experimental evidence." *Economics Letters* 157: 159–162.
- Lührmann, Melanie, Marta Serra-Garcia, and Joachim Winter. 2018. "The Impact of Financial Education on Adolescents' Intertemporal Choices." *American Economic Journal: Economic Policy* 10 (3): 309–32.
- Lundberg, Shelly, Jennifer L. Romich, and Kwok Ping Tsang. 2009. "Decision-Making by Children." *Review of Economics of the Household* 7 (1): 1–30.
- Mather, Nancy, and Richard W. Woodcock. 2001. *Woodcock-Johnson III Tests of Achievement: Examiners Manual*. Itasca, IL: Riverside Pub.
- Mayer, Susan E., Ariel Kalil, Philip Oreopoulos, and Sebastian Gallegos. 2019. "Using behavioral insights to increase parental engagement the parents and children together intervention." *Journal of Human Resources* 54 (4): 900–925.
- Meier, Stephan, and Charles D. Sprenger. 2010. "Present-Biased Preferences and Credit Card Borrowing." *American Economic Journal: Applied Economics* 2 (1): 193–210.
- Meier, Stephan, and Charles D. Sprenger. 2013. "Discounting Financial Literacy: Time Preferences and Participation in Financial Education Programs." *Journal of Economic Behavior & Organization* 95: 159–74.
- Meristo, Marek, and Luca Surian. 2013. "Do Infants Detect Indirect Reciprocity?" *Cognition* 129 (1): 102–113.
- Mischel, Walter, and Ebbe B. Ebbesen. 1970. "Attention in Delay of Gratification." *Journal of Personality and Social Psychology* 16 (2): 329–37.
- Mischel, Walter, Yuichi Shoda, and Monica I. Rodriguez. 1989. "Delay of Gratification in Children." *Science* 244 (4907): 933–38.
- Mischel, Walter, Yuichi Shoda, and Philip K. Peake. 1988. "The Nature of Adolescent Competencies Predicted by Preschool Delay of Gratification." *Journal of Personality and Social Psychology* 54 (4): 687.
- Moffitt, Terrie E., Louise Arseneault, Daniel Belsky, Nigel Dickson, Robert J. Hancox, HonaLee Harrington, Renate Houts, et al. 2011. "A Gradient of Childhood Self-Control Predicts Health, Wealth, and Public Safety." *Proceedings of the National Academy of Sciences* 108 (7): 2693–98.

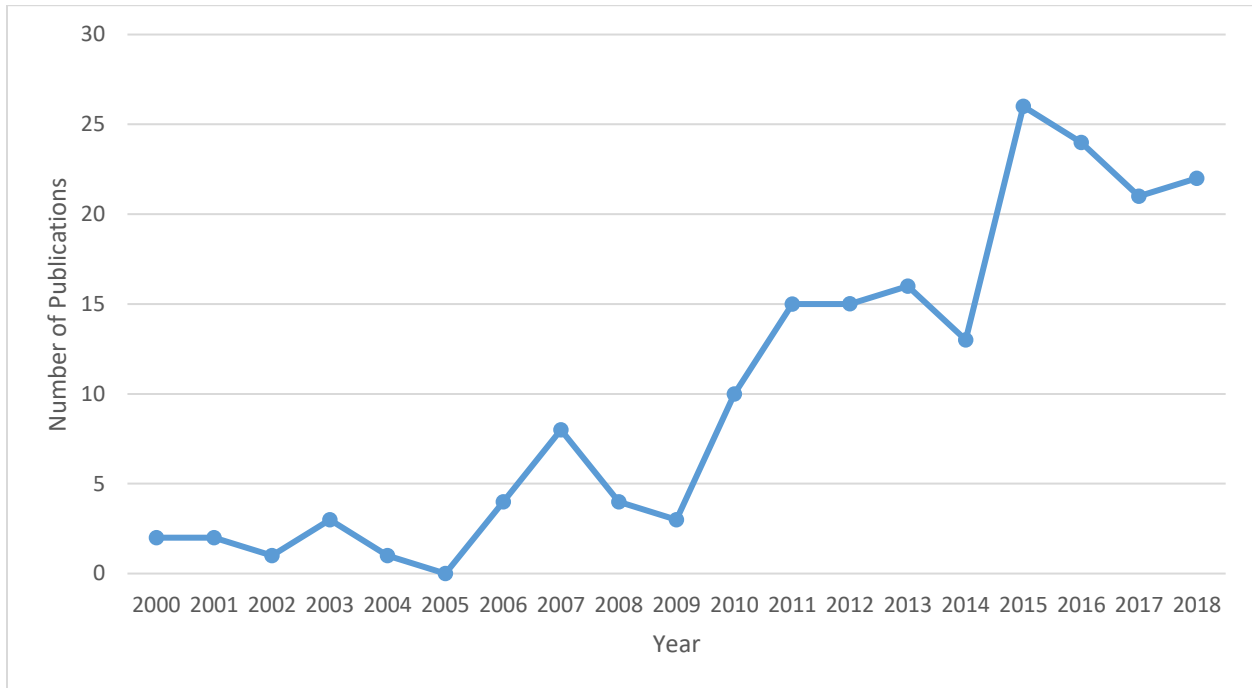
- Morgan, Myfanwy, Sara Gibbs, Krista Maxwell, and Nicky Britten. 2002. "Hearing Children's Voices: Methodological Issues in Conducting Focus Groups with Children Aged 7-11 Years." *Qualitative Research* 2 (1): 5–20.
- Morphett, Mabel Vogel, and Carleton Washburne. "When should children begin to read?." *The Elementary School Journal* 31, no. 7 (1931): 496-503.
- Naglieri, Jack A, and Sam Goldstein. 2014. *Handbook of Executive Functioning*. Springer.
- Niederle, Muriel, and Lise Vesterlund. 2007. "Do Women Shy Away From Competition? Do Men Compete Too Much?" *The Quarterly Journal of Economics* 122 (3): 1067–1101.
- Niederle, Muriel, and Lise Vesterlund. 2011. "Gender and Competition." *Annual Review of Economics* 3 (1): 601–30.
- Noble, Kimberly G., Bruce D. McCandliss, and Martha J. Farah. 2007. "Socioeconomic Gradients Predict Individual Differences in Neurocognitive Abilities." *Developmental Science* 10 (4): 464–80.
- Ólafsson, Kiartan, Sonia Livingstone, and Leslie Haddon. 2013. "How to Research Children and Online Technologies? Frequently Asked Questions and Best Practice." *How to Research Children and Online Technologies? Frequently Asked Questions and Best Practice*. London: EU Kids Online.
- Parker, Susan W., and Petra E. Todd. 2017. "Conditional Cash Transfers: The Case of Progreso/Oportunidades." *Journal of Economic Literature* 55 (3): 866-915.
- Petrie, Ragan, and Carmit Segal. 2017. "Gender Differences in Competitiveness: The Role of Prizes." *Working Paper*.
- Picone, Gabriel, Frank Sloan, and Donald Taylor. 2004. "Effects of Risk and Time Preference and Expected Longevity on Demand for Medical Tests." *Journal of Risk and Uncertainty* 28 (1): 39–53.
- Pischke, Jörn-Steffen, and Till von Wachter. 2008. "Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation." *Review of Economics and Statistics* 90 (3): 592–98.
- Potter, Mary C., and Ellen I. Levy. 1968. "Spatial Enumeration without Counting." *Child Development*.
- Pujol, Jesús, Pere Vendrell, Carme Junqué, Josep L. Martí-Vilalta, and Antoni Capdevila. 1993. "When Does Human Brain Development End? Evidence of Corpus Callosum Growth up to Adulthood." *Annals of Neurology* 34 (1): 71–75.
- Punch, Samantha. 2002. "Research with Children: The Same or Different from Research with Adults?" *Childhood* 9 (3): 321–41.
- Rao, Gautam. 2019. "Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools." *American Economic Review* 109 (3): 774–809.
- Rao, Uma, Tanuj Sidhartha, Karen R. Harker, Anup S. Bidesi, Li-Ann Chen, and Monique Ernst. "Relationship between adolescent risk preferences on a laboratory task and behavioral measures of risk-taking." *Journal of Adolescent Health* 48, no. 2 (2011): 151-158.
- Rea, David, and Tony Burton. "New evidence on the Heckman curve." *Journal of Economic Surveys* 34, no. 2 (2020): 241-262.
- Riach Peter A., Judy Rich. 2002. "Field experiments of discrimination in the market place" *The Economic Journal*, 112 (483) : F480-F518
- Rosenthal, Robert, and Lenore Jacobson. 1968. "Pygmalion in the Classroom." *The Urban Review* 3 (1): 16–20.

- Sadoff, Sally. 2014. "The Role of Experimentation in Education Policy." *Oxford Review of Economic Policy* 30 (4): 597–620.
- Samek, Anya, Andre Gray, Ashlesha Datar, and Nancy Nicosia. 2019. "Adolescent Time and Risk Preferences: Measurement, Determinants and Field Consequences." *CESR-Schaeffer Working Paper Forthcoming*.
- Schneider, Wolfgang. 2011. "Memory Development in Childhood." In *The Wiley-Blackwell Handbook of Childhood Cognitive Development*, 2nd ed., 347–76. Blackwell Publishing.
- Schwarz, J. Conrad, Janet B. Schrager, and Andrea E. Lyons. 1983. "Delay of Gratification by Preschoolers: Evidence for the Validity of the Choice Paradigm." *Child Development*, 620–25.
- Sher, Itai, Melissa Koenig, and Aldo Rustichini. 2014. "Children's Strategic Theory of Mind." *Proceedings of the National Academy of Sciences* 111 (37): 13307–12.
- Shloim, Netalie, Lisa R. Edelson, Nathalie Martin, and Marion M. Hetherington. 2015. "Parenting Styles, Feeding Styles, Feeding Practices, and Weight Status in 4–12 Year-Old Children: A Systematic Review of the Literature." *Frontiers in psychology* 6: 1849.
- Shoda, Yuichi, Walter Mischel, and Philip K. Peake. 1990. "Predicting Adolescent Cognitive and Self-Regulatory Competencies from Preschool Delay of Gratification: Identifying Diagnostic Conditions." *Developmental Psychology* 26 (6): 978.
- Smith, Vernon L. 1962. "An Experimental Study of Competitive Market Behavior." *Journal of Political Economy* 70 (2): 111–37.
- Smith, Vernon L. 1982. "Microeconomic Systems as an Experimental Science." *The American Economic Review* 72 (5): 923–955.
- Sparks, Erin, Meghan G. Schinkel, and Chris Moore. 2017. "Affiliation Affects Generosity in Young Children: The Roles of Minimal Group Membership and Shared Interests." *Journal of Experimental Child Psychology* 159: 242–62.
- Steinbeis, Nikolaus, and Eveline A. Crone. 2016. "The Link Between Cognitive Control and Decision-Making Across Child and Adolescent Development." *Current Opinion in Behavioral Sciences* 10: 28–32.
- Sternberg, Robert J. 2011. "Individual Differences in Cognitive Development." In *The Wiley-Blackwell Handbook of Childhood Cognitive Development*, 2nd ed., 749–74. Blackwell Publishing.
- Sutter, Matthias, and Daniela Glätzle-Rützler. 2015. "Gender Differences in the Willingness to Compete Emerge Early in Life and Persist." *Management Science* 61 (10): 2339–54.
- Sutter, Matthias, and Martin G. Kocher. 2007. "Trust and Trustworthiness across Different Age Groups." *Games and Economic Behavior* 59 (2): 364–82.
- Sutter, Matthias, Claudia Zoller, and Daniela Glätzle-Rützler. 2019. "Economic Behavior of Children and Adolescents – A First Survey of Experimental Economics Results." *European Economic Review* 111: 98–121.
- Sutter, Matthias, Martin G. Kocher, Daniela Glätzle-Rützler, and Stefan T. Trautmann. 2013. "Impatience and Uncertainty: Experimental Decisions Predict Adolescents Field Behavior." *American Economic Review* 103 (1): 510–31.
- Sylvia, Sean, Nele Warrinnier, Renfu Luo, Ai Yue, Orazio Attanasio, Alexis Medina, and Scott Rozelle. "From quantity to quality: Delivering a home-based parenting intervention through China's family planning cadres." *The Economic Journal* (2020).

- Tan, Lynne S. C., and Peter S. C. Bryant. 2000. "The Cues That Infants Use to Distinguish Discontinuous Quantities: Evidence Using a Shift-Rate Recovery Paradigm." *Child Development* 71 (5): 1162–78.
- Tanaka, Tomomi, Colin F Camerer, and Quang Nguyen. 2010. "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam." *American Economic Review* 100 (1): 557–71.
- Watts, Tyler W., Greg J. Duncan, and Haonan Quan. "Revisiting the marshmallow test: A conceptual replication investigating links between early delay of gratification and later outcomes." *Psychological science* 29, no. 7 (2018): 1159-1177.
- Weller, Rosalyn E., Edwin W. Cook, Kathy B. Avs-ar, and James E. Cox. 2008. "Obese Women Show Greater Delay Discounting than Healthy-Weight Women." *Appetite* 51 (3): 563–69.
- Wellman, Henry M, David Cross, and Julianne Watson. 2001. "Meta-Analysis of Theory-of-Mind Development: The Truth about False Belief." *Child Development* 72 (3): 655–84.
- Wilde, Louis. 1980. "On the Use of Laboratory Experiments in Economics." in Joseph Pitt, ed., *The Philosophy of Economics*, Dordrecht: Reidel.
- Williams, Janet. 1999. "Test Review." *Canadian Journal of School Psychology* 14 (2): 67–73.
- Wolf, Michael S, and Terry C Davis. 2004. "Health Literacy: Implications for Family Medicine." *FAMILY MEDICINE-KANSAS CITY* 36 (8): 595–98.
- Wolfers, Justin. 2006. "Diagnosing Discrimination: Stock Returns and Ceo Gender." *Journal of the European Economic Association* 4 (2-3): 531–41.
- Ye, Karen J., Anya Samek, Keith J. Yoder, Jean Decety, Ali Hortacsu, and John A. List. 2021. "Early Childhood Programs Change Test Scores but do They Change Brain Activity?" *Work in Progress*.
- Yeager, David S., Paul Hanselman, Gregory M. Walton, Jared S. Murray, Robert Crosnoe, Chandra Muller, Elizabeth Tipton et al. "A national experiment reveals where a growth mindset improves achievement." *Nature* 573, no. 7774 (2019): 364-369.
- Zhu, Yi, Xian Guan, and Yansong Li. 2015. "The Effects of Intergroup Competition on Prosocial Behaviors in Young Children: a Comparison of 2.5-3.5 Year-Olds with 5.5-6.5 Year-Olds." *Frontiers in Behavioral Neuroscience* 9: 16.
- Zur, Osnat, and Rochel Gelman. 2004. "Young Children Can Add and Subtract by Predicting and Checking." *Early Childhood Research Quarterly* 19 (1): 121–37.

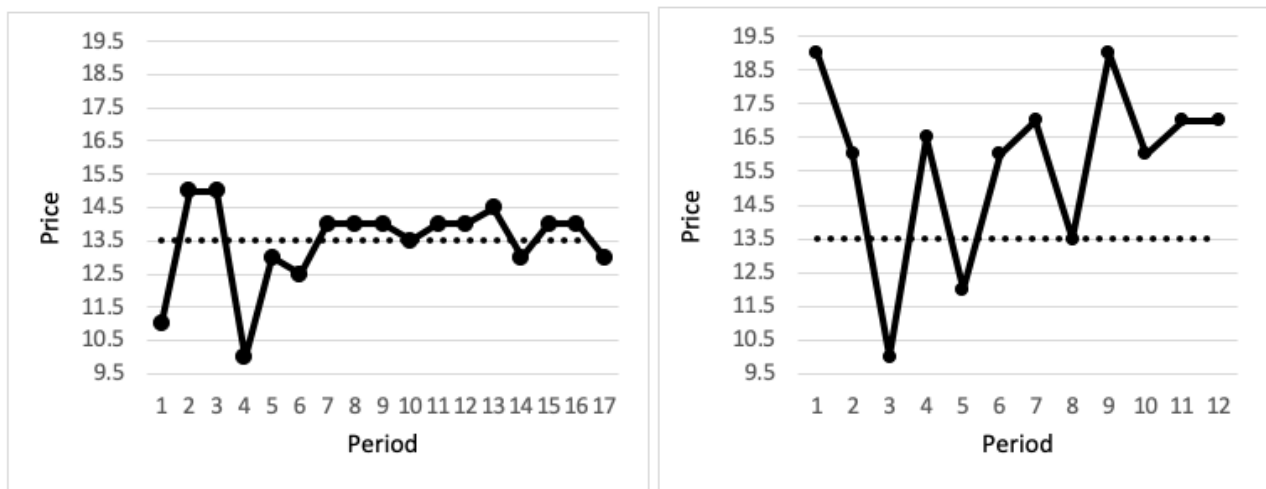
FIGURES AND TABLES

Figure 1: Growth in Economic Experiments with Children



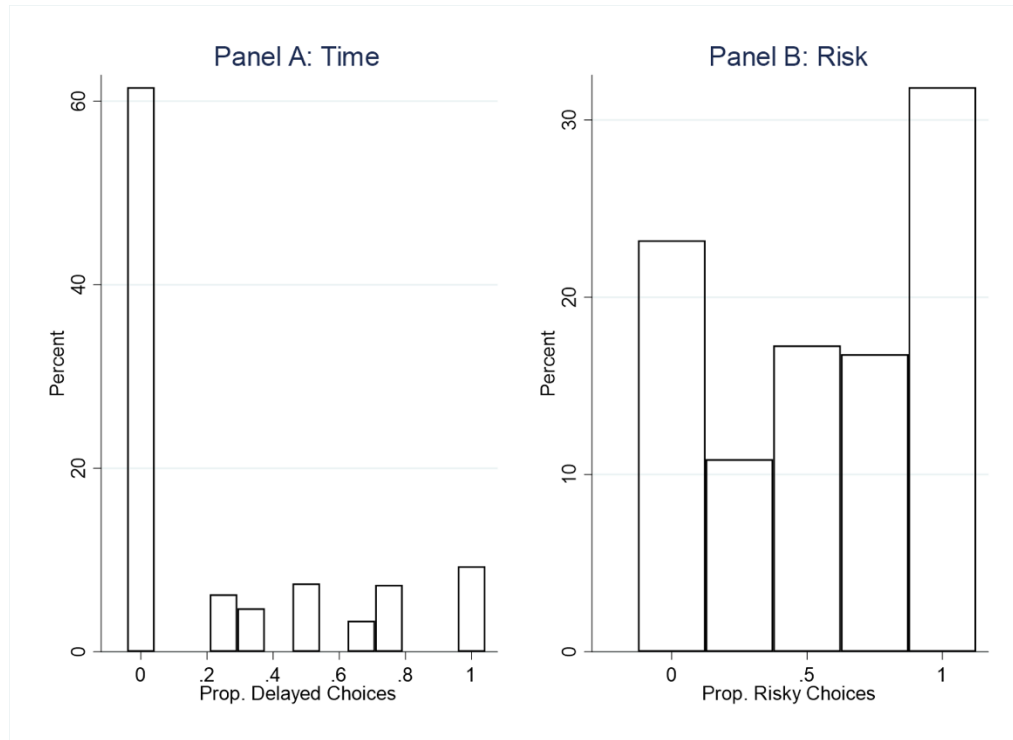
Notes: This figure shows the growth in publications (including working papers posted online) that use experiments with children, between 2000 and today. The papers included here come from a comprehensive search of the literature by 1) searching Google Scholar with keywords “economic experiment” (including searches specific to economic games commonly used in the laboratory) and “children” or “adolescents” and 2) reaching out to the Economic Science Association (ESA) mailing list and to authors of papers identified on Google Scholar to find papers that we may have missed. More details about our search are available in the appendix.

Figure 2: Markets with Children in List (2004a)



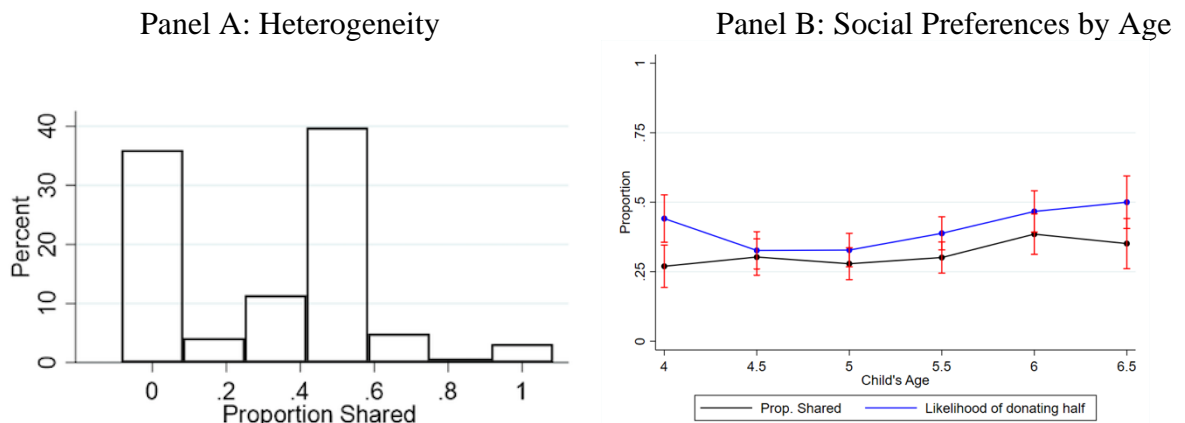
Notes: Reproduction of Figures 6 and 7 from List (2004a). Experienced children (left) and inexperienced children (right).

Figure 3: Heterogeneity in Time and Risk Preferences of Young Children



Notes: This figure uses data from CHECC reported on in Castillo et al. (2020) to plot a histogram showing heterogeneity in time preferences (Panel A) and risk preferences (Panel B). Data is depicted as the proportion of delayed choices in a menu of earlier/later decisions and proportion of risky choices in a menu of safe/risky decisions.

Figure 4: Heterogeneity in Social Preferences and Social Preferences by Age



Notes: This figure uses data from CHECC reported on in Castillo et al. (2020) to plot a histogram showing heterogeneity in social preferences (Panel A) and preferences by child age, from 4 to 6.6 years old (Panel B). Data is depicted as proportion of the reward shared (out of 6). Here, we only use “Wave 2012” data because this is the data for which we have the largest age range. Panel B also includes the proportion who share exactly half of their endowment.

Table 1: Aspects of Experiment Design by Child Development Age

	Preschool (3-5)	Early Elementary (6-8)	Late Elementary (9-11)	Early Adolescence (12-14)	Late Adolescence (15-17)	Standard Practices
Instruction Delivery Format One-on-one Group NP N	55.42% 42.17% 2.41% 83	42.28% 54.47% 3.25% 123	31.78% 66.36% 1.87% 107	25.81% 68.82% 5.38% 93	29.17% 66.67% 4.17% 72	Group
Identity of Partner/Group In-Person Non-Human Photo Gender Age Anonymous Other N	33.33% 2.47% 16.05% 14.81% 17.28% 20.99% 8.64% 81	24.63% 4.48% 13.43% 13.43% 23.13% 25.37% 10.45% 134	25.69% 4.59% 8.26% 10.09% 20.18% 33.03% 6.42% 109	30.95% 2.38% 1.19% 7.14% 16.67% 47.62% 4.76% 84	26.67% 1.67% 0% 5% 16.67% 51.67% 6.67% 60	Anonymous
Instructions Electronic Physical Written Oral NP N	2.41% 9.64% 0% 86.75% 1.20% 83	4.07% 4.88% 0.81% 88.62% 1.63% 123	3.74% 2.80% 1.87% 89.72% 1.87% 107	6.45% 3.23% 5.38% 81.72% 3.23% 93	15.28% 2.78% 6.94% 73.61% 1.39% 72	Experimenter reads instructions out loud
Instruction Aids Physical Graphical Videos None N	33.64% 15.45% 0.91% 50.00% 110	27.87% 17.49% 0.55% 54.10% 183	21.94% 16.13% 1.29% 60.65% 155	14.73% 24.81% 0% 60.47% 129	8.85% 25.66% 0% 65.49% 113	
Subject Decision Format Electronic Physical Written Oral NP N	5.45% 67.27% 5.45% 20.00% 1.82% 110	7.65% 45.36% 24.59% 20.77% 1.64% 183	11.61% 32.90% 36.77% 16.77% 1.94% 155	19.38% 26.36% 42.64% 10.08% 1.55% 129	23.89% 14.16% 53.98% 7.96% 0% 113	Electronic (e.g., zTree), rarely on paper
Comprehension Check Electronic Physical Written Oral None N	3.64% 14.55% 2.73% 40.91% 38.18% 110	2.19% 9.84% 3.28% 40.98% 43.72% 183	3.23% 4.52% 5.16% 35.48% 51.61% 155	2.33% 4.65% 7.75% 26.36% 58.91% 129	3.54% 1.77% 7.08% 16.81% 70.80% 113	Typically yes

	Preschool (3-5)	Early Elementary (6-8)	Late Elementary (9-11)	Early Adolescence (12-14)	Late Adolescence (15-17)	Standard Practices
Total Number of Subject Decisions						
Min	1	1	1	1	1	~20-40
Max	50	100	54	320	320	
Mean	5.964	5.615	5.904	12.747	13.357	
SD	8.909110	10.071	7.630	36.831	39.109	
N	110	182	154	129	108	
Experiment Length						
Min	11	3	3	5	5	~1-2 hours
Max	120	120	120	120	120	
Mean	35.475	34.164	41.007	49.25	47	
SD	29.248	27.581	30.401	33.237	34.755	
N	110	183	155	129	113	
Payout for Repeated Decisions						
All Rounds	87.04%	75.79%	63.33%	49.32%	31.25%	One randomly selected, or All
One Round	12.96%	24.21%	34.44%	45.21%	67.19%	
Multiple But Not All	0%	0%	2.22%	5.48%	1.56%	
N	110	95	90	73	64	
Incentives						
Token/Store	29.09%	42.62%	43.87%	25.58%	12.39%	Money (Cash)
Money		11.48%	24.52%	53.49%	76.11%	
Gift Card				5.43%	6.19%	
Candy	27.27%	18.03%	12.26%	6.98%		
Food			6.45%		2.65%	
Toys	10.00%					
Stickers	25.45%	12.02%	5.81%			
Stationery	8.18%	6.56%		3.88%	1.77%	
Other	0%	9.29%	7.10%	4.65%	0.88%	
N	110	183	155	129	113	
Exchange Rate						
1:1	55.26%	27.78%	14.94%	22.41%	13.16%	Varies
1:Many	15.79%	50%	63.22%	60.34%	68.42%	
NP	28.59%	22.22%	21.84%	17.24%	18.42%	
N	38	90	87	58	38	
Game Type						
Risk	10.00%	12.02%	13.55%	21.71%	29.20%	All
Time	10.91%	10.93%	12.90%	10.85%	15.04%	
Dictator	46.36%	37.16%	33.55%	26.36%	11.50%	
Strategy	13.64%	16.39%	15.48%	11.63%	7.96%	
Competition	3.64%	5.46%	6.45%	11.63%	11.50%	
Ultimatum	7.27%	8.74%	7.74%	8.53%	7.08%	
Other	8.18%	9.29%	10.32%	9.30%	7.96%	
N	110	183	155	129	113	

Notes: This tables summarizes the 190 papers found that used economics experiments with children. The appendix provides detail about how this information was summarized.

Appendix

A.1 Paper Identification Criteria

The papers were included if they met all of the following conditions:

1. They conduct an “economic game” using children under the age of 18 years,
2. They provide an outcome-based incentive to children participating,
3. They do not deceive the participants in any form,
4. And if the child has a “real” partner (when applicable).

The “economic games” that we report on are – all allocation (dictator and ultimatum), fairness, real effort, risk, strategy and rationality (coordination, competition, prisoner’s dilemma, etc.), time, and trust games.

Vast economic literature supports incentivizing task behavior to gauge actual behavior. Hence, we only reviewed papers where the outcome of the task the child participated in determined his/her payout.

It is standard practice in economics experiments to not deceive the subjects in any form. Consequently, we consider papers where the children were not deceived, and if they had a partner, the partner was a real person (e.g., not a puppet, stuffed toy, or hand-drawn picture of another child or hypothetical child).

In 2016, we sent emails asking for papers that met this criteria. The email was sent to the esa-discuss list and to researchers who were authors or appeared in citation lists of papers we had identified to that point.

We searched Google Scholar with relevant terms, such as “Preschoolers Dictator Task,” “Real Effort Experiment Children,” or “Coordination Task Adolescents,” to find additional papers that met our criteria. We completed a thorough search through 2018, and include papers from 2019 to the best of our ability.

Each paper was entered twice on Excel and compared by at least two different undergraduate research assistants to minimize errors. Two authors were emailed for additional clarification. In April, 2019, all unpublished papers were searched on Google Scholar for updated journal publication status.

A.2 Inclusion and Exclusion Criteria for Table 1

Table 1 presents a summary of 197 papers we identified.

We limit our findings to children over the age of 3 years.

In addition to how experiments were conducted with children, we present a column with standard practices as used in economics experiments with university students. These were compiled using the authors' knowledge of that literature and not based on a particular reference.

A.3 Description of the Development of Table 1

Next, we provide an explanation of the variables presented in Table 1. Our coding process is detailed along with a brief list of examples found in the literature. The section features an asterisk (*) if the coding followed a hierarchy system. By this we mean that a study used two or more of the formats and we chose the best term that encompassed the other responses. For example, for subject decision format, the hierarchy was electronic > physical > written > oral. This means that if the subject responded by announcing their answer verbally and also typing it in electronically it was coded as "electronic" only. We coded it this way because it would have been too difficult to try to explain the combination of formats used, and we felt it would not be as informative to use a "select all that apply" approach. The format types are listed in order of the hierarchy where applicable.

1. Instructions*: How instructions were delivered to the children. Possible media include –
 - a. Electronic – children received instructions on an electronic device, such as a vignette showcased on a tablet or computer
 - b. Physical – experimenters conducted a live-demonstration of the procedure for the children to follow along, or acted out a sequence to depict a plotline
 - c. Written – children were presented with a textual guide detailing their tasks, or experimenters wrote instructions on a board in a class
 - d. Oral – children followed along with verbal commands announced by the experimenter
 - e. NP – experimenters do not provide a method for how children were instructed

When instructions were given in multiple media, we used the hierarchy system (above) wherein the superior medium could include any of the subordinate media. E.g.: Instructions provided on an electronic tablet could be accompanied with verbal commands. Only one response in this category was chosen per experiment.

2. Instruction Aids: Were the children instructed using any additional tools. Possible tools include –
 - a. Physical – experimenters used actual objects and props, such as urns and marbles
 - b. Graphical – children were shown pictures or charts of supplemental materials
 - c. Videos – children viewed clips or vignettes
 - d. None – no additional facilitators used

Only one response in this category was chosen per experiment.

3. Instruction Delivery Format: How many children received instructions at the same time. Possible options include –
 - a. One-on-one – each child received instructions individually
 - b. Group – children received instructions in a cohort

- c. NP – experimenters do not provide information on session sizes

Only one response in this category was chosen per experiment.

- 4. Identity of Partner/Group*: Who the children played with or against. Partners could be identified using the possible options –
 - a. In-Person – child’s partner is physically present or they know the exact identity, such as when experimenters identify the child’s “mother” or “math teacher”
 - b. Non-Human – the partner is a computer or an institution, such as a charity or NGO
 - c. Photo – child is shown a picture of another child
 - d. Gender – child is told the gender of their partner, such as when experimenters tell the child their partner is “another boy/girl (gender matched)”
 - e. Age – partner’s age is revealed to the child, such as when experimenters tell a child their partner is “another child in their class”
 - f. Anonymous – the child is not given any identifying information about the partner
 - g. Other – the child is given any other information, such as the socio-economic status or lingual-group membership of their partner

This variable also followed a hierarchy system when children were given multiple identifiers. This hierarchy follows the same order as above, but with ties for In-Person and Non-Human, and Gender and Age. This was conditional on being a group task. Only one response in this category was chosen per experiment.

- 5. Subject Decision Format*: How the children made their choices. Possible media are –
 - a. Electronic – children responded by making decisions on an electronic device
 - b. Physical – the child had to operate his/her own body to perform their preferred response, such as pulling a lever to make an allocation decision
 - c. Written – the child had to mark or hand-write their response
 - d. Oral – the child made a verbal announcement as a response
 - e. NP – the experimenters do not provide information on how children made their decisions

When the children made their responses in multiple media, we coded it using a hierarchy system (above). Only one response in this category was chosen per experiment.

- 6. Total Number of Subject Decision: How many times did the child play that game. This count includes each decision round of a game, such as every row in a multiple price list (MPL). E.g.: A child plays three variations of a time-preferences elicitation task, with five decisions each, here the total number of subject decisions would be the product, 15. This was coded as a numeric response.
- 7. Experiment Length: How long on average did the total experimental procedure last per child. This is presented in minutes. When the publications specified a range, we include in our calculations the median time as most studies only provide the average length of gameplay. If the experimenters gave both the length of the gameplay and instructional

period, we take the combined total time for consistency, as other studies only provide the total experiment length. This was coded as a numeric response.

8. Payout for Repeated Decisions: How many of the rounds played per child were incentivized for a payout. Number of incentivized rounds were coded as follows –
- All Rounds – each round of the game affected the child’s total reward, such as in a gambling task
 - One Round – only one of the multiple rounds of the game count for the outcome, such as in a risk-lottery task
 - Multiple But Not All – some of the multiple rounds affect the child’s reward

This was conditional on children receiving a reward in the game played. Only one response in this category was chosen per experiment.

9. Incentives: What the children were offered as the outcome of the game or compensation. Additional incentives examples, segregated by age, are presented in Appendix A.4. Prizes were coded as one of the following –
- Token/Store – when a child is offered any non-monetary object that holds a fixed intrinsic value, as determined by the experimenter, which can be later exchanged for a desirable good. The store could include candies, snack foods, toys, clips and bracelets, stickers, or coloring and drawing supplies.
 - Money – when children were given a cash reward or tokens that would later be translated to money, usually in the local currency
 - Gift Card – children’s rewards were translated and loaded on a gift card, usually from a locally recognized vendor
 - Candy – children received sweets as prizes, such as Skittles, M&Ms, or regionally-preferred sweet-snacks
 - Food – children received other edible foods as prizes, such as regionally-preferred food-snacks like maize, or bags of chips/crisps and wafers.
 - Toys – children received age-appropriate toys as prizes, such as stuffed animals, toy cars, and clay dough
 - Stickers – children received individual or sticker packs as reward, such as animal or cartoon character stickers, shiny/glittering and scented stickers, and large-sized stickers
 - Stationery – children received writing or drawing supplies, such as crayons, color pencils, erasers, and pens

All applicable responses in this category were chosen.

10. Exchange Rate: How the child’s experimental outcome was converted to a reward. This was coded as one of the following –
- 1:1 – each experimental token/currency unit was exchanged for one reward
 - 1:Many – the child received multiple rewards per unit they earned
 - NP – the experimenters do not provide information on how the reward were translated from the experimental procedure

Only one response in this category was chosen per experiment.

11. Game Type: What games were played with children. Each publication could include multiple games. The game types were coded as follows –
- Risk – children played risk-preferences elicitation games, such as a gambling task or a lottery task
 - Time – children played a time-preferences elicitation game, such as the marshmallow experiment or delayed rewards
 - Dictator – children played the dictator task, or a variation of the game
 - Strategy – children played a coordination game, prisoner’s dilemma game, investment game, or any variation of such games
 - Competition – children played a competitive game against another child
 - Ultimatum – children played the ultimatum task with another child
 - Other – children played any other game, such as a real effort, fairness, trust, public goods, market preferences, cheating game, or any such variation

Only one response in this category was chosen per experiment.

12. Comprehension Check*: How the experimenters’ gauged the child’s understanding of the instructions provided. Their questions were framed as follows –
- Electronic – children answered checks using an electronic device, such as a tablet or computer
 - Physical – children had to perform a sample or a trial of the actual game
 - Written – children marked or wrote their response to signify understanding
 - Oral – children announced their verbal response to questions from an experimenter
 - None – no indication of whether children were tested on their understanding of the instructions; this was also used when this information was not provided in the experiment procedure or within the child instruction script

If the checks were conducted in multiple media, we follow the hierarchy system (above). Only one response in this category was chosen per experiment.

13. Sample size: This is the number of children in the analysis sample of the papers. When the experiments in the paper were comprised of multiple age groups and the breakdown of children by group was not specified, we assumed there were an equal number of children from each age group. We only include children who were part of the analysis sample (e.g., children who participated in a pilot that was later not evaluated are not included). This was coded as a numeric response.
14. N: This is the number of studies reviewed that included this age group and had information about the experimental aspect. This was coded as a numeric response.
15. Incentives: The most common incentives for each age group are given at the bottom of Table A1. This was coded as a string response.

Table A1: Further Detail from Table 1

Experimental Aspect \ Developmental Stage (Ages)	Under 3 (Ages 0-2)	Preschool (Ages 3-5)	Early Elementary (6-8)	Late Elementary (9-11)	Early Adolescence (12-14)	Late Adolescence (15-17)
Continents						
Africa	0%	2.41%	2.44%	2.80%	5.38%	6.94%
Asia	50%	20.48%	21.14%	24.30%	16.13%	12.50%
Europe	16.67%	31.33%	42.28%	52.34%	47.31%	54.17%
North America	33.33%	48.19%	36.59%	29.91%	36.56%	30.56%
South America	0%	4.82%	4.88%	5.61%	2.15%	2.78%
Australia	0%	1.20%	0%	0.93%	2.15%	2.78%
Sample Size						
Min	18	9	9	10	10	11
Max	157	360	1919	3000	1413	2983
Mean	72.333	106.205	190.667	240.720	212.172	240.083
SD	48.343	3.942	294.741	378.723	267.181	405.640
N	6	83	123	107	93	72
Year						
Min	2011	2001	1966	1966	1966	1966
Max	2018	2019	2019	2019	2019	2019
Mean	2014.167	2013.410	2012.943	2012.720	2012.505	2012.903
SD	2.639	3.942	6.402	6.599	6.499	7.053
N	6	83	123	107	93	72
Incentives						
Sweets	Marshmallows Chocolate-Bars	Skittles M&Ms Gummy-Bears Chocolate-Bars	Skittles M&Ms Chocolate-Bars	Skittles M&Ms Chocolate-Bars	Skittles M&Ms Chocolate-Bars	Skittles M&Ms Chocolate-Bars
Stationery	Stickers	Stickers Color-Pencils Wax-Crayons	Scented-Stickers Pens Pencils Erasers Art - Supplies	Pens Pencils Erasers Art & School Supplies	Pens Pencils Erasers Markers Highlighters Notebooks & Notepads	Pens Pencils Markers Highlighters Notebooks & Notepads
Toys	Stuffed-Animals	Plastic-Animals Balloons Bouncy-Balls	Bracelets Silly Bandz Hair-Clips Miniature-vehicles	Miniature-vehicles Silly Bandz Beads	Silly Bandz Bouncy-Balls	
Food	Cookies	Cookies Wafers Fruits Ice-Cream	Cookies Fruits Juice-Box Ice-Cream	Fruits Juice-Box Chips Bags Ice-Cream	Fruits Juice-Box Chips Bags Ice-Cream	Fruits Juice-Box Chips Bags

A.5 Inclusion and Exclusion Criteria for Figure 1

Figure 1 presents the year of publication for 190 studies that met our criteria.

We include studies conducted with children of all ages (0-17).

We limit our findings to years 2000-2018.

A.6 Reference List of all Papers Included In Table 1

- Aguilar-Pardo, David, Rosario Martínez-Arias, and Fernando Colmenares. 2013. "The Role of Inhibition in Young Children's Altruistic Behaviour." *Cognitive Processing* 14 (3): 301–7.
- Alan, Sule, and Seda Ertac. 2015. "Patience, Self-Control and the Demand for Commitment: Evidence from a Large-Scale Field Experiment." *Journal of Economic Behavior & Organization* 115: 111–22.
- Alan, Sule, Nazli Baydar, Teodora Boneva, Thomas F. Crossley, and Seda Ertac. 2017. "Transmission of Risk Preferences from Mothers to Daughters." *Journal of Economic Behavior & Organization* 134: 60–77.
- Alan, Sule, and Seda Ertac. 2018. "Mitigating the Gender Gap in the Willingness to Compete: Evidence from a Randomized Field Experiment." *Journal of the European Economic Association* 17 (4): 1147–85.
- Alan, Sule, and Seda Ertac. 2018. "Fostering Patience in the Classroom: Results from Randomized Educational Intervention." *Journal of Political Economy* 126 (5): 1865–1911.
- Alan, Sule, Teodora Boneva, and Seda Ertac. 2019. "Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit." *The Quarterly Journal of Economics* 134 (3): 1121–62.
- Alencar, Al, José de Oliveira Siqueira, and Maria Emilia Yamamoto. 2008. "Does Group Size Matter? Cheating and Cooperation in Brazilian School Children☆." *Evolution and Human Behavior* 29 (1): 42–48.
- Almås, Ingvid, Alexander W. Cappelen, Erik Ø. Sørensen, and Bertil Tungodden. 2010. "Fairness and the Development of Inequality Acceptance." *Science* 328 (5982): 1176–78.
- Almås, Ingvid, Alexander W. Cappelen, Kjell G. Salvanes, Erik Ø. Sørensen, and Bertil Tungodden. 2016. "Fairness and Family Background." *Politics, Philosophy & Economics* 16 (2): 117–31.
- Andersen, Steffen, Seda Ertac, Uri Gneezy, John A. List, and Sandra Maximiano. 2013. "Gender, Competitiveness, and Socialization at a Young Age: Evidence From a Matrilineal and a Patriarchal Society." *Review of Economics and Statistics* 95 (4): 1438–43.
- Andreoni, James, Michael Kuhn, John A. List, and Anya Samek. 2017. "Field Experiments on the Development of Time Preferences." *The Field Experiments Website*, no. 00615.
- Andreoni, James, Amalia Di Girolamo, John A. List, Claire Mackevicius, and Anya Samek. 2019. "Risk Preferences of Children and Adolescents in Relation to Gender, Cognitive Skills, Soft Skills, and Executive Functions." *Journal of Economic Behavior & Organization*.
- Angerer, Silvia, Daniela Glätzle-Rützler, Philipp Lergetporer, and Matthias Sutter. 2015. "Donations, Risk Attitudes and Time Preferences: A Study on Altruism in Primary School Children." *Journal of Economic Behavior & Organization* 115: 67–74.

- Angerer, Silvia, Philipp Lergetporer, Daniela Glätzle-Rützler, and Matthias Sutter. 2015. "How to Measure Time Preferences in Children: a Comparison of Two Methods." *Journal of the Economic Science Association* 1 (2): 158–69. doi:10.1007/s40881-015-0016-0.
- Angerer, Silvia, Daniela Glätzle-Rützler, Philipp Lergetporer, and Matthias Sutter. 2016. "Cooperation and Discrimination within and across Language Borders: Evidence from Children in a Bilingual City." *European Economic Review* 90: 254–64.
- Angerer, Silvia, E. Glenn Dutcher, Daniela Glätzle-Rützler, Philipp Lergetporer, and Matthias Sutter. 2017. "Gender Differences in Discrimination Emerge Early in Life: Evidence from Primary School Children in a Bilingual City." *Economics Letters* 152: 15–18.
- Antrobus, Emma, Victoria Baranov, Deborah A. Cobb-Clark, Lorraine Mazerolle, and Anieszka Tymula. 2017. "The Risk and Time Preferences of Young Truants and Their Parents."
- Apicella, Coren L., and Anna Dreber. 2014. "Sex Differences in Competitiveness: Hunter-Gatherer Women and Girls Compete Less in Gender-Neutral and Male-Centric Tasks." *Adaptive Human Behavior and Physiology* 1 (3): 247–69.
- Armin, Falk, and Fabian Kosse. 2016. "Early Childhood Environment, Breastfeeding and the Formation of Preferences." *SSRN Electronic Journal*.
- Apestequia, Jose, Steffen Huck, Jörg Oechssler, Elke Weidenholzer, and Simon Weidenholzer. 2018. "Imitation of Peers in Children and Adults." *Games* 9 (1): 11.
- Avinun, Reut, et al. 2011. "AVPR1A Variant Associated with Preschoolers Lower Altruistic Behavior." *PLoS ONE* 6 (9).
- Barash, Jori, Isabelle Brocas, Juan D. Carrillo, and Niree Kodaverdian. 2019. "Heuristic to Bayesian: The Evolution of Reasoning from Childhood to Adulthood." *Journal of Economic Behavior & Organization* 159: 305–22.
- Bartling, Björn, Ernst Fehr, and Daniel Schunk. 2012. "Health Effects on Children's Willingness to Compete." *Experimental Economics* 15 (1): 58–70.
- Bauer, Michal, Julie Chytilová, and Barbara Pertold-Gebicka. 2013. "Parental Background and Other-Regarding Preferences in Children." *Experimental Economics* 17 (1): 24–46.
- Bašić, Zvonimir, Armin Falk, and Fabian Kosse. 2019. "The Development of Egalitarian Norm Enforcement in Childhood and Adolescence." *Journal of Economic Behavior & Organization*.
- Bekkers, Rene. 2017. "Do Two Eyes See More Generosity Than One?"
- Belot, Michèle, and Jeroen Van De Ven. 2011. "Friendships and Favouritism on the Schoolground – A Framed Field Experiment." *The Economic Journal* 121 (557): 1228–51.
- Benenson, Joyce F., Joanna Pascoe, and Nicola Radmore. 2007. "Children's Altruistic Behavior in the Dictator Game." *Evolution and Human Behavior* 28 (3): 168–75.
- Ben-Ner, Avner, John A. List, Louis Putterman, and Anya Samek. 2017. "Learned Generosity? An Artefactual Field Experiment with Parents and Their Children." *Journal of Economic Behavior & Organization* 143: 28–44.
- Benzio, Avi, and Gil Diesendruck. 2015. "Parochialism in Preschool Boys Resource Allocation." *Evolution and Human Behavior* 36 (4): 256–64.
- Bereby-Meyer, Yoella, and Shelly Fiks. 2013. "Changes in Negative Reciprocity as a Function of Age." *Journal of Behavioral Decision Making* 26 (4): 397–403.
- Bettinger, Eric, and Robert Slonim. 2007. "Patience among Children." *Journal of Public Economics* 91 (1-2): 343–63.

- Bindra, Parampreet Christopher, Daniela Glätzle-Rützler, and Philipp Lergetporer. 2018. "Discrimination at Young Age: Experimental Evidence from Preschool Children." *CESifo Working Paper*, no. 7396.
- Blake, Peter R., and David G. Rand. 2010. "Currency Value Moderates Equity Preference among Young Children." *Evolution and Human Behavior* 31 (3): 210–18.
- Blake, Peter R., and Katherine McAuliffe. 2011. "'I Had so Much It Didn't Seem Fair': Eight-Year-Olds Reject Two Forms of Inequity." *Cognition* 120 (2): 215–24.
- Blake, Peter R., Marco Piovesan, Natalia Montinari, Felix Warneken, and Francesca Gino. 2015. "Prosocial Norms in the Classroom: The Role of Self-Regulation in Following Norms of Giving." *Journal of Economic Behavior & Organization* 115: 18–29.
- Blake, Peter R., et al. 2015. "The Ontogeny of Fairness in Seven Societies." *Nature* 528 (7581): 258–61.
- Blake, Peter R., David G. Rand, Dustin Tingley, and Felix Warneken. 2015. "The Shadow of the Future Promotes Cooperation in a Repeated Prisoner's Dilemma for Children." *Scientific Reports* 5 (1).
- Blake, Peter R., John Corbit, Tara C. Callaghan, and Felix Warneken. 2016. "Give as I Give: Adult Influence on Children's Giving in Two Cultures." *Journal of Experimental Child Psychology* 152: 149–60.
- Booth, Alison L., and Patrick Nolen. 2012. "Choosing to Compete: How Different Are Girls and Boys?" *Journal of Economic Behavior & Organization* 81 (2): 542–55.
- Booth, Alison L., and Patrick Nolen. 2012. "Gender Differences in Risk Behaviour: Does Nurture Matter?" *The Economic Journal* 122 (558).
- Borghans, Lex, Bart H.H. Golsteyn, James Heckman, and Huub Meijers. 2009. "Gender Differences in Risk Aversion and Ambiguity Aversion." *Journal of the European Economic Association* 7 (2-3): 649–58.
- Brocas, Isabelle, Juan D. Carrillo, and Niree Kodaverdian. 2017. "Altruism and Strategic Giving in Children and Adolescent." *CEPR Discussion Paper*, no. DP12288.
- Brocas, Isabelle, and Juan D. Carrillo. 2018. "Iterative Dominance in Young Children: Experimental Evidence in Simple Two-Person Games." *Journal of Economic Behavior & Organization*.
- Brocas, Isabelle, Juan D. Carrillo, T. Dalton Combs, and Niree Kodaverdian. 2019. "The Development of Consistent Decision-Making across Economic Domains." *Games and Economic Behavior* 116: 217–40.
- Brosig-Koch, Jeannette, Timo Heinrich, and Christoph Helbach. 2015. "Exploring the Capability to Reason Backwards: An Experimental Study with Children, Adolescents, and Young Adults." *European Economic Review* 74: 286–302.
- Buccioli, Alessandro, and Marco Piovesan. 2011. "Luck or Cheating? A Field Experiment on Honesty with Children." *Journal of Economic Psychology* 32 (1): 73–78.
- Buccioli, Alessandro, Daniel Houser, and Marco Piovesan. 2011. "Temptation and Productivity: A Field Experiment with Children." *Journal of Economic Behavior & Organization* 78 (1-2): 126–36.
- Bueno-Guerra, Nereida, David Leiva, Montserrat Colell, and Josep Call. 2016. "Do Sex and Age Affect Strategic Behavior and Inequity Aversion in Children?" *Journal of Experimental Child Psychology* 150: 285–300.
- Bunch, Katie M., Glenda Andrews, and Graeme S. Halford. 2007. "Complexity Effects on the Childrens Gambling Task." *Cognitive Development* 22 (3): 376–83.

- Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. 2014. "Gender, Competitiveness and Career Choices." *The Quarterly Journal of Economics* 129 (3): 1409–47.
- Bügelmayer, Elisabeth, and C. Katharina Spiess. 2014. "Spite and Cognitive Skills in Preschoolers." *Journal of Economic Psychology* 45: 154–67.
- Cappelen, Alexander W., John A. List, Anya Samek, and Bertil Tungodden. 2019. "The Effect of Early Childhood Education on Social Preferences." *Journal of Political Economy*.
- Castelli, Ilaria, Davide Massaro, Alan G. Sanfey, and Antonella Marchetti. 2010. "Fairness and Intentionality in Children's Decision-Making." *International Review of Economics* 57 (3): 269–88.
- Castillo, Marco, Paul J. Ferraro, Jeffrey L. Jordan, and Ragan Petrie. 2011. "The Today and Tomorrow of Kids: Time Preferences and Educational Outcomes of Children." *Journal of Public Economics* 95 (11-12): 1377–85.
- Castillo, Marco. 2020. "Negative Childhood Experiences and Risk Aversion: Evidence from Children Exposed to Domestic Violence." *IZA DP No. 13320*.
- Castillo, Marco, Jeffrey L. Jordan, and Ragan Petrie. 2018. "Children's Rationality, Risk Attitudes and Field Behavior." *European Economic Review* 102: 62–81.
- Castillo, Marco, Jeffrey L. Jordan, and Ragan Petrie. 2018. "Discount Rates of Children and High School Graduation." *The Economic Journal* 129 (619): 1153–81.
- Chen, Jingnan, Daniel Houser, Natalia Montinari, and Marco Piovesan. 2016. "Beware of Popular Kids Bearing Gifts: A Framed Field Experiment." *Journal of Economic Behavior & Organization* 132: 104–20.
- Chiang, Yen-Sheng, and Chyi-In Wu. 2015. "Social Influence and the Adaptation of Parochial Altruism: a Dictator-Game Experiment on Children and Adolescents under Peer Influence." *Evolution and Human Behavior* 36 (6): 430–37.
- Chowdhury, Shyamal, Matthias Sutter, and Klaus Zimmermann. 2018. "Evaluating Intergenerational Persistence of Economic Preferences: A Large Scale Experiment with Families in Bangladesh." *SSRN Electronic Journal*.
- Chung, Dongil, Kyongsik Yun, Jin Ho Kim, Bosun Jang, and Jaeseung Jeong. 2011. "Different Gain/Loss Sensitivity and Social Adaptation Ability in Gifted Adolescents during a Public Goods Game." *PLoS ONE* 6 (2).
- Cipriani, Marco, Paola Giuliano, and Olivier Jeanne. 2013. "Like Mother like Son? Experimental Evidence on the Transmission of Values from Parents to Children." *Journal of Economic Behavior & Organization* 90: 100–111.
- Cohn, Alain, and Michel André Maréchal. 2018. "Laboratory Measure of Cheating Predicts School Misconduct." *The Economic Journal* 128 (615): 2743–54.
- Corbit, John, Katherine McAuliffe, Tara C. Callaghan, Peter R. Blake, and Felix Warneken. 2017. "Children's Collaboration Induces Fairness Rather than Generosity." *Cognition* 168: 344–56.
- Cotton, Christopher, Frank McIntyre, and Joseph Price. 2013. "Gender Differences in Repeated Competition: Evidence from School Math Contests." *Journal of Economic Behavior & Organization* 86: 52–66.
- Cowell, Jason M., Anya Samek, John A. List, and Jean Decety. 2015. "The Curious Relation between Theory of Mind and Sharing in Preschool Age Children." *Plos One* 10 (2).
- Cox, James, John A. List, Michael Price, Vjollca Sadiraj, and Anya Samek. 2016. "Moral Costs and Rational Choice: Theory and Experimental Evidence." *Experimental Economics*.

- Czermak, Simon, Francesco Feri, Daniela Glätzle-Rützler, and Matthias Sutter. 2016. "How Strategic Are Children and Adolescents? Experimental Evidence from Normal-Form Games." *Journal of Economic Behavior & Organization* 128: 265–85.
- Cárdenas, Juan-Camilo, Anna Dreber, Emma Von Essen, and Eva Ranehill. 2015. "Cooperativeness and Competitiveness in Children." *Journal of Behavioral and Experimental Economics* 59: 32–41.
- Cárdenas, Juan-Camilo, Anna Dreber, Emma Von Essen, and Eva Ranehill. 2012. "Gender Differences in Competitiveness and Risk Taking: Comparing Children in Colombia and Sweden." *Journal of Economic Behavior & Organization* 83 (1): 11–23.
- Cárdenas, Juan-Camilo, Anna Dreber, Emma Von Essen, and Eva Ranehill. 2014. "Gender and Cooperation in Children: Experiments in Colombia and Sweden." *PLoS ONE* 9 (3).
- Deckers, Thomas, Armin Falk, Fabian Kosse, and Hannah Schildberg-Hörisch. 2015. "How Does Socio-Economic Status Shape a Child's Personality?" *IZA Discussion Paper Series*, no. 8977.
- Detlefsen, Lena, Andreas Friedl, Katharina Lima De Miranda, Ulrich Schmidt, and Matthias Sutter. 2018. "Are Economic Preferences Shaped by the Family Context? The Impact of Birth Order and Siblings' Sex Composition on Economic Preferences." *SSRN Electronic Journal*.
- Dreber, Anna, Emma Von Essen, and Eva Ranehill. 2013. "Gender and Competition in Adolescence: Task Matters." *Experimental Economics* 17 (1): 154–72.
- Eckel, Catherine C., Philip J. Grossman, Cathleen A. Johnson, Angela C.M. De Oliveira, Christian Rojas, and Rick K. Wilson. 2010. "(Im)Patience Among Adolescents: A Methodological Note." *SSRN Electronic Journal*.
- Eckel, Catherine C., Philip J. Grossman, Cathleen A. Johnson, Angela C.M. De Oliveira, Christian Rojas, and Rick K. Wilson. 2011. "Social Norms of Sharing in High School: Teen Giving in the Dictator Game." *Journal of Economic Behavior & Organization* 80 (3): 603–12.
- Eckel, Catherine C., Philip J. Grossman, Cathleen A. Johnson, Angela C.M. De Oliveira, Christian Rojas, and Rick K. Wilson. 2012. "School Environment and Risk Preferences: Experimental Evidence." *Journal of Risk and Uncertainty* 45 (3): 265–92.
- Falk, Armin, Fabian Kosse, Pia Rosina Dovert-Pinger, Hannah Schildberg-Hörisch, and Thomas Deckers. 2019. "Socio-Economic Status and Inequalities in Children's IQ and Economic Preferences." *Working Paper*, no. 166.
- Fan, Chinn-Ping. 2000. "Teaching Children Cooperation — An Application of Experimental Game Theory." *Journal of Economic Behavior & Organization* 41 (3): 191–209.
- Fehr, Ernst, Helen Bernhard, and Bettina Rockenbach. 2008. "Egalitarianism in Young Children." *Nature* 454 (7208): 1079–83.
- Fehr, Ernst, Daniela Glätzle-Rützler, and Matthias Sutter. 2013. "The Development of Egalitarianism, Altruism, Spite and Parochialism in Childhood and Adolescence." *European Economic Review* 64: 369–83.
- Felfer, Christina, Martin G. Kocher, Helmut Rainer, Judith Saurer, and Thomas Siedler. 2018. "More Opportunity, More Cooperation? The Behavioral Effects of Birthright Citizenship on Immigrant Youth."
- Garon, Nancy, Brittany Johnson, and Ashley Steeves. 2011. "Sharing with Others and Delaying for the Future in Preschoolers." *Cognitive Development* 26 (4): 383–96.

- Garon, Nancy M., Julie Longard, Susan E. Bryson, and Chris Moore. 2012. "Making Decisions about Now and Later: Development of Future-Oriented Self-Control." *Cognitive Development* 27 (3): 314–22.
- Geng, Sen, Yujia Peng, Jason Shachat, and Huizhen Zhong. 2015. "Adolescents, Cognitive Ability, and Minimax Play." *Economics Letters* 128: 54–58.
- Glätzle-Rützler, Daniela, Manuela Oberauer, Matthias Sutter, and Levent Yilmaz. 2016. "How Children and Adults Play Experimental Games with Equilibria in Mixed Strategies."
- Glätzle-Rützler, Daniela, Matthias Sutter, and Achim Zeileis. 2015. "No Myopic Loss Aversion in Adolescents? – An Experimental Note." *Journal of Economic Behavior & Organization* 111: 169–76.
- Goeree, Jacob K., Margaret A. McConell, Tiffany Mitchell, Tracey Tromp, and Leeat Yariv. 2010. "The 1/d Law of Giving." *American Economic Association* 2 (1): 183–203.
- Goldvicht-Bacon, Efrat, and Gil Diesendruck. 2016. "Children's Capacity to Use Cultural Focal Points in Coordination Problems." *Cognition* 149: 95–103.
- Grueneisen, Sebastian, Emily Wyman, and Michael Tomasello. 2014. "'I Know You Don't Know I Know...' Children Use Second-Order False-Belief Reasoning for Peer Coordination." *Child Development* 86 (1): 287–93.
- Grueneisen, Sebastian, Emily Wyman, and Michael Tomasello. 2014. "Children Use Salience to Solve Coordination Problems." *Developmental Science* 18 (3): 495–501.
- Guinote, Ana, Ioanna Cotzia, Sanpreet Sandhu, and Pramila Siwa. 2015. "Social Status Modulates Prosocial Behavior and Egalitarianism in Preschool Children and Adults." *Proceedings of the National Academy of Sciences* 112 (3): 731–36.
- Gummerum, Michaela, Monika Keller, Masanori Takezawa, and Jutta Mata. 2008. "To Give or Not to Give: Children's and Adolescents' Sharing and Moral Negotiations in Economic Decision Situations." *Child Development* 79 (3): 562–76.
- Gummerum, Michaela, Yaniv Hanoch, Monika Keller, Katie Parsons, and Alegra Hummel. 2010. "Preschoolers' Allocations in the Dictator Game: The Role of Moral Emotions." *Journal of Economic Psychology* 31 (1): 25–34.
- Hamann, Katharina, Felix Warneken, Julia R. Greenberg, and Michael Tomasello. 2011. "Collaboration Encourages Equal Sharing in Children but Not in Chimpanzees." *Nature* 476 (7360): 328–31.
- Hamann, Katharina, Johanna Bender, and Michael Tomasello. 2014. "Meritocratic Sharing Is Based on Collaboration in 3-Year-Olds." *Developmental Psychology* 50 (1): 121–28.
- Hao, Jian. 2017. "Do Children with Better Inhibitory Control Donate More? Differentiating between Early and Middle Childhood and Cool and Hot Inhibitory Control." *Frontiers in Psychology* 8.
- Harbaugh, William T. 2000. "Children's Altruism in Public Good and Dictator Experiments." *Economic Inquiry* 38 (1): 95–109.
- Harbaugh, William T., Kate S. Krause, and Lise Vesterlund. 2001. "Are Adults Better Behaved than Children? Age, Experience, and the Endowment Effect." *Economics Letters* 70 (2): 175–81.
- Harbaugh, William T., Kate S. Krause, and Timothy R. Berry. 2001. "GARP for Kids: On the Development of Rational Choice Behavior." *American Economic Review* 91 (5): 1539–45.
- Harbaugh, William T., Kate S. Krause, and Lise Vesterlund. 2002. "Risk Attitudes of Children and Adults: Choices over Small and Large Probability Gains and Losses." *Experimental Economics* 5 (1): 53–84.

- Harbaugh, William T., Kate S. Krause, and Steven G. Liday. 2003. "Bargaining by Children." *SSRN Electronic Journal*.
- Harbaugh, William T., Kate S. Krause, Steven G. Liday, and Lise Vesterlund. 2003. "Trust in Children." *Trust and Reciprocity: Interdisciplinary Lessons from Experimental Research*, 302–22.
- Harbaugh, William T., Kate S. Krause, and Lise Vesterlund. 2007. "Learning to Bargain." *Journal of Economic Psychology* 28 (1): 127–42.
- Heinrich, Timo, and Jason Shachat. 2018. "The Development of Risk Aversion and Prudence in Chinese Children and Adolescents."
- Hermes, Henning, Florian Hett, Mario Mechtel, Felix Schmidt, Daniel Schunk, and Valentin Wagner. 2019. "Do Children Cooperate Conditionally? Adapting the Strategy Method for First-Graders." *Journal of Economic Behavior & Organization*.
- Hoffmann, Robert, and Jin-Yee Tee. 2006. "Adolescent–Adult Interactions and Culture in the Ultimatum Game." *Journal of Economic Psychology* 27 (1): 98–116.
- House, Bailey, Joseph Henrich, Barbara Sarnecka, and Joan B. Silk. 2013. "The Development of Contingent Reciprocity in Children." *Evolution and Human Behavior* 34 (2): 86–93.
- Houser, Daniel, and Daniel Schunk. 2009. "Social Environments with Competitive Pressure: Gender Effects in the Decisions of German Schoolchildren." *Journal of Economic Psychology* 30 (4): 634–41.
- Houser, Daniel, Natalia Montinari, and Marco Piovesan. 2012. "Private and Public Decisions in Social Dilemmas: Evidence from Children's Behavior." *PLoS ONE* 7 (8).
- Hu, Yu, and Yi Zhu. 2018. "Exploring an Age Difference in Preschool Children's Competitiveness Following a Competition." *Frontiers in Psychology* 9.
- Häger, Kirsten, Bastian Oud, and Daniel Schunk. 2012. "Egalitarian Envy: Cross-Cultural Variation in the Development of Envy in Children." *Jena Economic Research Papers*, no. 059.
- Jennings, Amanda Brooke. 2016. "Are You More Economic than a First Grader?: A Mixed Methods Approach in a Common Pool Resources Experiment."
- John, Katrin, and Stephan L. Thomsen. 2015. "School-Track Environment or Endowment: What Determines Different Other-Regarding Behavior across Peer Groups?" *Games and Economic Behavior* 94: 122–41.
- John, Katrin, and Stephan L. Thomsen. 2017. "Gender Differences in the Development of Other-Regarding Preferences."
- Khachatryan, Karen, Anna Dreber, Emma Von Essen, and Eva Ranehill. 2015. "Gender and Preferences at a Young Age: Evidence from Armenia." *Journal of Economic Behavior & Organization* 118: 318–32.
- Khadjavi, Menusch, and Andreas Nicklisch. 2018. "Parents' Ambitions and Children's Competitiveness." *Journal of Economic Psychology* 67: 87–102.
- Kogut, Tehila. 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): 226–36.
- Kogut, Tehila, Paul Slovic, and Daniel Västfjäll. 2015. "The Effect of Recipient Identifiability and Neediness on Children's Sharing Behavior." *Journal of Behavioral Decision Making* 29 (4): 353–62.

- Kogut, Tehila, and Paul Slovic. 2016. "The Development of Scope Insensitivity in Sharing Behavior." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 42 (12): 1972–81.
- Kosse, Fabian, and Friedhelm Pfeiffer. 2012. "Impatience among Preschool Children and Their Mothers." *Economics Letters* 115 (3): 493–95.
- Kosse, Fabian, Thomas Deckers, Pia Pinger, Hannah Schildberg-Hörisch, and Armin Falk. 2019. "The Formation of Prosociality: Causal Evidence on the Role of Social Environment." *Journal of Political Economy*.
- Kramer, Julia Lucia, Silvia Lucia Lübbecke, and Nina Lucia Stephan. 2019. "Do Children Stop Cheating When Someone Else Reaps the Benefits? A Lab in the Field Experiment." *Paderborn University, Faculty of Business Administration and Economics*, no. 33.
- Lemmon, Karen, and Chris Moore. 2007. "The Development of Prudence in the Face of Varying Future Rewards." *Developmental Science* 10 (4): 502–11.
- Lergetporer, Philipp, Silvia Angerer, Daniela Glätzle-Rützler, and Matthias Sutter. 2014. "Third-Party Punishment Increases Cooperation in Children through (Misaligned) Expectations and Conditional Cooperation." *Proceedings of the National Academy of Sciences* 111 (19): 6916–21.
- Levin, Irwin P., and Stephanie S. Hart. 2003. "Risk Preferences in Young Children: Early Evidence of Individual Differences in Reaction to Potential Gains and Losses." *Journal of Behavioral Decision Making* 16 (5): 397–413.
- Li, Rosa, Rachel C. Roberts, Scott A. Huettel, and Elizabeth M. Brannon. 2017. "Five-Year-Olds Do Not Show Ambiguity Aversion in a Risk and Ambiguity Task with Physical Objects." *Journal of Experimental Child Psychology* 159: 319–26.
- Lima de Miranda, Katharina. 2019. "Mindfulness, Preferences and Well-Being: Mindfulness Predicts Adolescents' Field Behavior." *Kiel Working Paper*, no. 1227.
- List, John A., and Daniel L. Millimet. 2008. "The Market: Catalyst for Rationality and Filter of Irrationality." *The B.E. Journal of Economic Analysis & Policy* 8 (1).
- List, Annika J., John A. List, and Anya Samek. 2017. "Discrimination among Pre-School Children: Field Experimental Evidence." *Economics Letters* 157: 159–62.
- List, John A., and Anya C. Samak. 2013. "Exploring the Origins of Charitable Acts: Evidence from an Artefactual Field Experiment with Young Children." *Economics Letters* 118 (3): 431–34.
- Liu, Buyun, et al. 2016. "Altruistic Sharing Behavior in Children: Role of Theory of Mind and Inhibitory Control." *Journal of Experimental Child Psychology* 141: 222–28.
- Lombardi, Elisabetta, Cinzia Di Dio, Ilaria Castelli, Davide Massaro, and Antonella Marchetti. 2017. "Prospective Thinking and Decision Making in Primary School Age Children." *Heliyon* 3 (6).
- Lührmann, Melanie, Konstantin Lucks, and Joachim K. Winter. 2017. "Peer Effects in Risky Choices among Adolescents."
- Lührmann, Melanie, Marta Serra-Garcia, and Joachim Winter. 2018. "The Impact of Financial Education on Adolescents' Intertemporal Choices." *American Economic Journal: Economic Policy* 10 (3): 309–32.
- Maggian, Valeria, and Marie Claire Villeval. 2015. "Social Preferences and Lying Aversion in Children." *Experimental Economics* 19 (3): 663–85.

- Majolo, Bonaventura, and Laëtitia Maréchal. 2017. "Between-Group Competition Elicits within-Group Cooperation in Children." *Scientific Reports* 7 (1).
- Marchetti, Antonella, et al. 2018. "Social Decision Making in Adolescents and Young Adults: Evidence From the Ultimatum Game and Cognitive Biases." *Psychological Reports* 122 (1): 135–54.
- Maria Giovanna, Devetag, Sibilla Di Guida, and Natalia Montinari. 2013. "Theory of Mind in the Social Sciences: An Experiment On Strategic Thinking in Children," 67–92.
- Martinsson, Peter, Katarina Nordblom, Daniela Glätzle-Rützler, and Matthias Sutter. 2011. "Social Preferences during Childhood and the Role of Gender and Age — An Experiment in Austria and Sweden." *Economics Letters* 110 (3): 248–51.
- McAuliffe, Katherine, Peter R. Blake, Grace Kim, Richard W. Wrangham, and Felix Warneken. 2013. "Social Influences on Inequity Aversion in Children." *PLoS ONE* 8 (12).
- McAuliffe, Katherine, Peter R. Blake, and Felix Warneken. 2014. "Children Reject Inequity out of Spite." *Biology Letters* 10 (12): 20140743.
- McAuliffe, Katherine, and Yarrow Dunham. 2017. "Fairness Overrides Group Bias in Children's Second-Party Punishment." *Journal of Experimental Psychology: General* 146 (4): 485–94.
- McGuigan, Nicola, Ruth Fisher, and Rory Glasgow. 2016. "The Influence of Receiver Status on Donor Prosociality in 6- to 11-Year-Old Children." *Child Development* 87 (3): 855–69.
- Meuwese, Rosa, Eveline A. Crone, Mark De Rooij, and Berna Güroğlu. 2014. "Development of Equity Preferences in Boys and Girls Across Adolescence." *Child Development* 86 (1): 145–58.
- Molino, José Alberto, Alfredo Ferrer, José Ignacio Giménez, Carlos Gracia-Lazaro, Yamir Moreno, and Angel Sanchez. 2016. "The Effect of Kinship on Intergenerational Cooperation: A Lab Experiment with Three Generations."
- Moreira, Bruno, Raul Matsushita, and Sergio Da Silva. 2010. "Risk Seeking Behavior of Preschool Children in a Gambling Task." *Journal of Economic Psychology* 31 (5): 794–801.
- Munro, Alistair, and Yuki Tanaka. 2014. "Risky Rotten Kids: an Experiment on amongst Adolescents in Rural Uganda." *National Graduate Institute for Policy Studies*, no. 14-01.
- Ongley, Sophia F., and Tina Malti. 2014. "The Role of Moral Emotions in the Development of Children's Sharing Behavior." *Developmental Psychology* 50 (4): 1148–59.
- Ouazad, Amine, and Lionel Page. 2013. "Students Perceptions of Teacher Biases: Experimental Economics in Schools." *Journal of Public Economics* 105: 116–30.
- Page, Lionel, Dipanwita Sarkar, and Juliana Silva-Goncalves. 2017. "The Older the Bolder: Does Relative Age among Peers Influence Children's Preference for Competition?" *Journal of Economic Psychology* 63: 43–81.
- Pappert, Anna-Theresa, Amanda Williams, and Chris Moore. 2016. "The Influence of Competition on Resource Allocation in Preschool Children." *Social Development* 26 (2): 367–81.
- Paulsen, David J., Michael L. Platt, Scott A. Huettel, and Elizabeth M. Brannon. 2011. "Decision-Making Under Risk in Children, Adolescents, and Young Adults." *Frontiers in Psychology* 2.
- Paulus, Markus, and Chris Moore. 2016. "Preschoolers' Generosity Increases with Understanding of the Affective Benefits of Sharing." *Developmental Science* 20 (3).

- Peters, H. Elizabeth, A. Sinan Unur, Jeremy Clark, and William D. Schulze. 2004. "Free-Riding and the Provision of Public Goods in the Family: A Laboratory Experiment." *International Economic Review* 45 (1): 283–99.
- Posid, Tasha, Allyse Fazio, and Sara Cordes. 2015. "Being Sticker Rich: Numerical Context Influences Children's Sharing Behavior." *PLoS One* 10 (11).
- Rao, Gautam. 2019. "Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools." *American Economic Review* 109 (3): 774–809.
- Rao, Uma, Tanuj Sidhartha, Karen R. Harker, Anup S. Bidesi, Li-Ann Chen, and Monique Ernst. 2011. "Relationship Between Adolescent Risk Preferences on a Laboratory Task and Behavioral Measures of Risk-Taking." *Journal of Adolescent Health* 48 (2): 151–58.
- Ruggeri, Azzurra, Shenghua Luan, Monika Keller, and Michaela Gummerum. 2017. "The Influence of Adult and Peer Role Models on Children' and Adolescents' Sharing Decisions." *Child Development* 89 (5): 1589–98.
- Sally, David, and Elisabeth Hill. 2006. "The Development of Interpersonal Strategy: Autism, Theory-of-Mind, Cooperation and Fairness." *Journal of Economic Psychology* 27 (1): 73–97.
- Samak, Anya C. 2013. "Is There a Gender Gap in Preschoolers' Competitiveness? An Experiment in the U.S." *Journal of Economic Behavior & Organization* 92: 22–31.
- Samek, Anya, Andre Gray, Ashlesha Datar, and Nancy Nicosia. "Adolescent Time and Risk Preferences: Measurement, Determinants and Field Consequences." *CESR-Schaeffer Working Paper* (2019).
- Schug, Joanna, Haruto Takagishi, Catalina Benech, and Hiroyuki Okada. 2016. "The Development of Theory of Mind and Positive and Negative Reciprocity in Preschool Children." *Frontiers in Psychology* 7.
- Sebastián-Enesco, Carla, María Victoria Hernández-Lloreda, and Fernando Colmenares. 2013. "Two and a Half-Year-Old Children Are Prosocial Even When Their Partners Are Not." *Journal of Experimental Child Psychology* 116 (2): 186–98.
- Shaw, Alex, Natalia Montinari, Marco Piovesan, Kristina R. Olson, Francesca Gino, and Michael I. Norton. 2014. "Children Develop a Veil of Fairness." *Journal of Experimental Psychology: General* 143 (1): 363–75.
- Silva, Phiética Raíssa Rodrigues Da, Natalia Andrea Cracciun Boccardi, Natalia Bezerra Dutra, Wallisen Tadashi Hattori, Maria Emília Yamamoto, and Anuska Irene Alencar. 2016. "Stickers versus Wafers: The Value of Resource in a Public Goods Game with Children." *Estudos De Psicologia* 21.
- Slovic, Paul. 1966. "Risk-Taking in Children: Age and Sex Differences." *Child Development* 37 (1): 169.
- Smith, Craig E., Peter R. Blake, and Paul L. Harris. 2013. "I Should but I Won't: Why Young Children Endorse Norms of Fair Sharing but Do Not Follow Them." *PLoS ONE* 8 (3).
- Sparks, Erin, Meghan G. Schinkel, and Chris Moore. 2017. "Affiliation Affects Generosity in Young Children: The Roles of Minimal Group Membership and Shared Interests." *Journal of Experimental Child Psychology* 159: 242–62.
- Sutter, Matthias. 2007. "Outcomes versus Intentions: On the Nature of Fair Behavior and Its Development with Age." *Journal of Economic Psychology* 28 (1): 69–78.
- Sutter, Matthias, and Martin G. Kocher. 2007. "Trust and Trustworthiness across Different Age Groups." *Games and Economic Behavior* 59 (2): 364–82.
- Sutter, Matthias, and Daniela Glätzle-Rützler. 2010. "Gender Differences in Competition Emerge Early in Life."

- Sutter, Matthias, Martin G. Kocher, Daniela Glätzle-Rützler, and Stefan T. Trautmann. 2013. "Impatience and Uncertainty: Experimental Decisions Predict Adolescents Field Behavior." *American Economic Review* 103 (1): 510–31.
- Sutter, Matthias, Levent Yilmaz, and Manuela Oberauer. 2015. "Delay of Gratification and the Role of Defaults—An Experiment with Kindergarten Children." *Economics Letters* 137: 21–24.
- Sutter, Matthias, and Daniela Glätzle-Rützler. 2015. "Gender Differences in the Willingness to Compete Emerge Early in Life and Persist." *Management Science* 61 (10): 2339–54.
- Sutter, Matthias, Daniela Glätzle-Rützler, Loukas Balafoutas, and Simon Czermak. 2016. "Cancelling out Early Age Gender Differences in Competition: an Analysis of Policy Interventions." *Experimental Economics* 19 (2): 412–32.
- Sutter, Matthias, Anna Untertrifaller, and Claudia Zoller. 2018. "Busy Little Bees—An Experiment on Diligence and Endogenous Time Scheduling in Early Childhood."
- Sutter, Matthias, Silvia Angerer, Daniela Glätzle-Rützler, and Philipp Lergetporer. 2018. "Language Group Differences in Time Preferences: Evidence from Primary School Children in a Bilingual City." *European Economic Review* 106: 21–34.
- Sutter, Matthias, Francesco Feri, Daniela Glätzle-Rützler, Martin G. Kocher, Peter Martinsson, and Katarina Nordblom. 2018. "Social Preferences in Childhood and Adolescence. A Large-Scale Experiment to Estimate Primary and Secondary Motivations." *Journal of Economic Behavior & Organization* 146: 16–30.
- Takagishi, Haruto, Shinya Kameshima, Joanna Schug, Michiko Koizumi, and Toshio Yamagishi. 2010. "Theory of Mind Enhances Preference for Fairness." *Journal of Experimental Child Psychology* 105 (1-2): 130–37.
- Takezawa, Masanori, Michaela Gummerum, and Monika Keller. 2006. "A Stage for the Rational Tail of the Emotional Dog: Roles of Moral Reasoning in Group Decision Making." *Journal of Economic Psychology* 27 (1): 117–39.
- Tymula, Anieszka, et al. 2012. "Adolescents Risk-Taking Behavior Is Driven by Tolerance to Ambiguity." *Proceedings of the National Academy of Sciences* 109 (42): 17135–40.
- Unur, A. Sinan, Elizabeth Peters, Kent D. Messer, and William D. Schulze. 2013. "Incentive Effects of Parents' Transfers to Children: An Artefactual Field Experiment." *The B.E. Journal of Economic Analysis & Policy* 13 (1): 73–106.
- Vaish, Amrisha, Esther Herrmann, Christiane Markmann, and Michael Tomasello. 2016. "Preschoolers Value Those Who Sanction Non-Cooperators." *Cognition* 153: 43–51.
- Vogelsang, Martina, Keith Jensen, Sebastian Kirschner, Claudio Tennie, and Michael Tomasello. 2014. "Preschoolers Are Sensitive to Free Riding in a Public Goods Game." *Frontiers in Psychology* 5.
- Weidenholzer, Elke. 2012. "Minimum Effort Game Experiment with Children."
- Weltzien, Sandra, Lauren E. Marsh, and Bruce Hood. 2018. "Thinking of Me: Self-Focus Reduces Sharing and Helping in Seven- to Eight-Year-Olds." *Plos One* 13 (1).
- Wyman, Emily, Hannes Rakoczy, and Mike Tomasello. 2013. "Non-Verbal Communication Enables Children's Coordination in a "Stag Hunt" Game." *European Journal of Developmental Psychology* 10 (5): 597–610.
- Yun, Kyongsik, Dongil Chung, Bosun Jang, Jin Ho Kim, and Jaeseung Jeong. 2011. "Mathematically Gifted Adolescents Have Deficiencies in Social Valuation and Mentalization." *PLoS ONE* 6 (4).

- Zhang, Jane. 2011. "Do Girls in China Compete Just As Much As Boys? Evidence from an Experiment That Predicts Educational Choice." *Technical Report*.
- Zhu, Yi, Xian Guan, and Yansong Li. 2015. "The Effects of Intergroup Competition on Prosocial Behaviors in Young Children: a Comparison of 2.5-3.5 Year-Olds with 5.5-6.5 Year-Olds." *Frontiers in Behavioral Neuroscience* 9 (December).
- Zhu, Yi, Jiajie Wang, Xiaofeng Lv, and Yansong Li. 2016. "Once Failed, Twice Shy: How Group-Based Competition Influences Risk Preference in Young Children." *Journal of Experimental Psychology: General* 145 (4): 397–401.