Does Preschool Affect Time Preferences? Evidence from a Field Experiment

Amanda Chuan, Anya Samek, Shreemayi Samujjwala

Michigan State University
University of California, San Diego
University of Pennsylvania

Last Modified: 8/26/2022

Abstract

We evaluate the causal impact of preschool on time preferences using a natural field experiment. After children finished preschool and entered elementary school, they participated in a 2-week experiment in which they chose between earlier and later rewards. Using these decisions, we examine patience, time inconsistency, and the demand for commitment. We find that children assigned to the preschool curriculum emphasizing self-regulation skills exhibited greater stability of patience relative to the control group. There is some evidence that control children were more likely than preschool children to change their decisions based on peer behavior, which may explain their greater impatience. Second, strategic use of the commitment device was exhibited by children in the preschool curriculum which focused on self-regulation, but not by other children. Third, we do not find treatment effects on time inconsistency. Lastly, we find that these behaviors positively predict math, reading, working memory, and attention up to five years later.

Acknowledgements: We thank John List and the team at the Chicago Heights Early Childhood Center (CHECC), especially Edie Dobrez and Kristin Troutman, for making this work possible. We also thank Charles Sprenger for invaluable input on the experimental design. This research was carried out while Samek was a visiting professor at the University of Chicago and was funded in part by the University of Chicago departmental research funds, the Kenneth and Anne Griffin Foundation and NIH grant 1R01DK114238. This research is approved by the Institutional Review Boards (IRB) at the University of Chicago, the University of Southern California and the University of California-San Diego.

1. Introduction

While the traditional view of preschool focuses on teaching the ABCs and the 123s, a growing literature has identified the importance of preschool in shaping non-cognitive skills (see Heckman, 2008; Carneiro and Heckman, 2003; Heckman, Stixrud, and Urzua, 2006; Fryer et al., 2015; Cappelen et al., 2020; Castillo et al., 2020; Nores et al., 2005; Belfield et al., 2006; Heckman et al., 2010; Campbell et al., 2001; Campbell et al., 2002; Masse and Barnett, 2002). In response, there are now a few preschool curricula focused on improving such skills (see Jones and Bouffard, 2012). For example, the *Tools of the Mind* curriculum aims to develop self-regulation and executive functioning through activities like Play Planning, which encourages children to develop independent plans for their activities each day (Bodrova and Leong, 2007). Prior experimental studies find that *Tools of the Mind* improves inhibitory control, working memory, cognitive flexibility, reasoning ability, reading, and math from preschool into first grade (Diamond et al., 2007; Diamond et al., 2019; Blair and Raver, 2014). More generally, preschool programs have been shown to improve cognitive ability, academic achievement, emotional development, and socioemotional regulation (Hahn et al., 2016; Ramon et al., 2018).

Another key non-cognitive skill that preschool may influence is time preferences. In particular, the activities that children are asked to do in preschool - e.g., sitting still for story time or taking turns to play with toys - may build patience, or the willingness to delay rewards. Further, Play Planning and related activities in programs like *Tools of the Mind* give children practice with setting intentions, making plans, and following through on these plans, potentially building self-control. This may in turn reduce time inconsistency, a feature of some time preference models that explains why people tend to make patient plans but fail to follow through in their actions (Strotz,

1955; Thaler and Shefrin, 1981; Herrnstein et al., 1993; Laibson, 1997; O'Donoghue and Rabin, 1999, O'Donoghue and Rabin, 2006).

In this paper, we evaluate the impact of preschool on time preferences. Our study has three key features. First, when children are 3-4 years old, they enroll in the Chicago Heights Early Childhood (CHECC) study and are randomized to a preschool or to a control group for 1-2 years. Within the preschool group, children are assigned to three different curricula that vary in their focus on self-regulation and executive function skills, including the *Tools of the Mind* curriculum, which has the greatest focus on executive functions; *Literacy Express & Math*, which focuses most on early literacy and math skills; and *CogX,* which combines instruction on literacy, math, and executive function skills. When children are 5-10 years old, we elicit their time preferences over a two-week period. The causal nature of the data is unique because most prior research on time preferences in children is correlational (List et al., 2021).

The second feature of our experiment is that it measures multiple constructs underlying time preferences. Patience and time consistency play necessary yet distinct roles in children's developmental trajectories.[1] A child may demonstrate patience by preferring a high test score in the future over the earlier reward of playing. However, she cannot accumulate skills through patience alone. She must also exhibit time consistency by resisting present-bias and repeatedly choosing to study even when playing provides immediate gratification. For this reason, our experiment focuses on the *stability of patience*, to examine whether preschool affects the persistence with which children exercise patience over a two-week period.

[1] More broadly, different measures of time preferences may measure different aspects of behavior. For example, Castillo et al. (2020) show that patience as measured by an incentivized economics experiment loads on a different factor compared to measures of self-control and delay-of-gratification from psychology.

The third key feature of our experiment is that we correlate early childhood time preferences to schooling outcomes of the children in middle childhood. There is some evidence that early childhood time preferences are associated with outcomes in middle childhood and adulthood (Castillo et al. 2020; Mischel et al. 1972, 1989; Shoda et al., 1990). However, the extent to which different aspects of time preferences matter is an open question. To answer these questions, we collect administrative data on standardized test scores, executive function measures, and disciplinary referrals up to five years following our experiment.

In our experiment, children are asked to choose between rewards over a two week period, following the designs of Augenblick et al. (2015), Augenblick and Rabin (2019), and Sadoff et al. (2020). On Day 1 of the first week, children choose between a smaller, earlier reward on Day 2, or a larger, later reward on Day 3. We refer to this as the "Advance Decision." We return to the children on Day 2 and surprise them with the opportunity to change their mind. Therefore, on Day 2 children choose again between the smaller, earlier reward on the same day, or a later, larger reward on the next day. We refer to this as the "Immediate Decision." We identify patience by the willingness to choose the larger, later reward in the Advance Decision, when neither option is paid immediately and children are not susceptible to present-bias. We identify time inconsistency based on whether children switch their choice between the Advance and Immediate Decisions. Children behave impulsively if they chose the later reward in the Advance Decision and switch to the earlier reward the next day in the Immediate Decision.

The second week is designed to evaluate the extent to which time preferences are stable across weeks and to assess the demand for commitment. Children again make Advance and Immediate Decisions. We refer to children who select the same option in the Advance Decision in both Week 1 and Week 2 as "stable", and children who select different options as "unstable." After

making their Advance Decision on Day 1, children are asked whether they would like to commit to their Advance Decision and restrict themselves from choosing the next day, when their future selves may be tempted by the earlier option that would be paid out immediately.

We report four sets of results. The first pertains to patience and time inconsistency in the overall sample. We find that 81% of children exhibit patience in Week 1 by choosing the later option. This proportion drops to 76% in Week 2, a significant difference of 5 percentage points. 21% of children exhibit inconsistency in Week 1 where their Advance Decision differed from their Immediate Decision. 74% of these inconsistent children were impulsive, choosing the delayed option at first but switching the next day when the earlier option would be paid out immediately. In Week 2, inconsistency drops to 19%, with 55% of inconsistent children displaying impulsivity.

Second, children assigned to *Tools of the Mind* tend to be patient in both weeks, whereas children in the other groups become less patient over time. Hence, while both the *Tools of the Mind* group and its Control group choose the delayed option 88% of the time in the Advance Decision in Week 1, there is a statistically significant difference in the choice of the delayed option in the Advance Decision by Week 2: 97% for *Tools of the Mind* and 74% for Control. For the *Literacy Express* and *CogX* groups, patience rates do not significantly differ from their respective Control groups in Week 1 or Week 2.

We evaluate several explanations for the difference in stability of patience. We find little evidence that it is due to learning from own behavior - patience in Week 2 has surprisingly little to do with patience and inconsistency in Week 1, which do not differ across groups. Rather, it appears that Control children are more susceptible to environmental influences than *Tools of the Mind* children. Control children are more likely to choose the earlier reward in Week 2 when more of their classmates ended up with the earlier reward in Week 1. In contrast, *Tools* children chose

the later reward in Week 2 independent of what their classmates did in Week 1. This contributes to the decline in patience among the Control group but not the *Tools* group. Susceptibility to peer influence does not significantly differ between the Control group and the other two Preschool curricula, *Literacy* and *CogX*. However, the point estimate for susceptibility among the *CogX* group falls between those for the Control and *Tools* groups, which may reflect the fact that the *CogX* curriculum was a hybrid of the *Tools of the Mind* and *Literacy Express* curricula.

Third, the *Tools* group is the only group that strategically uses commitment. In the overall sample, about half of children demand commitment, but commitment is uncorrelated with prior time inconsistency. The exception are *Tools* children, who are more likely to use commitment if they behaved impulsively in the previous week. Further evidence of strategic use is that *Tools* children who failed to hold out for the later option in the Immediate Decision were more likely to commit one week later compared to *Tools* children who successfully waited for the later option.

Our last set of results concern the relationship between time preferences and later outcomes in middle childhood. We find that patience at ages 5-10 are correlated with academic test scores and executive function measures at ages 7-12. Taken together, our results suggest some malleability of time preferences in early childhood and highlight the importance of time preferences in explaining academic and behavioral outcomes in middle childhood.

Our first contribution to the literature is to document the role of early education on time preferences. Several studies have considered whether targeted interventions change the time preferences of children and adolescents. Among adolescents, Lührmann et al. (2018) find that a financial literacy program reduces time inconsistency, and Alan and Ertac (2018) find that a program that teaches adolescents to imagine their future selves increases patience. Among young children, teaching waiting strategies improves delay of gratification (Shoda et al.,1990; Alan and

Ertac, 2014; 2015; Alan et al., 2021). However, Falk et al. (2020) note a need for research that "does not teach waiting strategies but provides an enhanced social environment" which can improve delay of gratification among children "deprived by the accident of birth." We find support for this recommendation in that peer influence is a powerful predictor of patience. This suggests that time preferences, like many other preferences, form in a social environment. Our preschool interventions provide a holistic learning environment over 1-2 years, and may have more lasting effects than shorter interventions that narrowly teach waiting strategies.

Second, we provide a deeper examination of the constructs underlying time preferences that are associated with academic and behavioral outcomes. As such, we join work by Castillo et al. (2018), who evaluate how adolescents' time preferences relate with disciplinary referrals in the next year; Andreoni et al. (2019), who evaluate the effect of CHECC on patience; and Castillo et al. (2020), who use the CHECC sample to evaluate the association of economic preferences at ages 3-5 with schooling outcomes at ages 6-12. Importantly, Andreoni et al. (2019) and Castillo et al. (2020) use one-shot time preference measures collected at the beginning of the CHECC study, whereas we take repeated time preference measures over two weeks. Similar to Andreoni et al. (2019), we do not find an effect of treatment in Week 1. However, we observe a time trend in our experiment that a one-shot experiment could not capture. We are better able to capture the stability of time preferences at ages 5-10, and to relate these unique measures to academic and behavioral outcomes at ages 5-15.[2] Our approach is more appropriate for examining human capital formation

[2] Other papers that use the CHECC study have explored the impact of programs on cognitive ability and executive function (Fryer et al., 2015; 2018), fairness preferences (Cappelen et al., 2020), cheating behavior (Houser et al., 2016), charitable giving (Ben-Ner et al., 2017; Samek and Sheremeta, 2017; List and Samak, 2013; List et al., 2018; Cowell et al., 2015; Cox et al., 2016), competitiveness (Samak, 2013), and food choice (Sadoff and Samek, 2019).

over childhood, since the stability of time preferences is more relevant for the pursuit of long-term goals than one-shot choices between earlier and later rewards.

Our third contribution is to policy regarding socioeconomic inequality. Our result that preschool increases patience highlights the value of preschool in improving noncognitive skills. This informs the debate about universal preschool in the United States, since children who are more patient have improved schooling outcomes, are more likely to graduate high school, and have better health outcomes (see Castillo et al., 2018; 2020). Moreover, our sample consists mainly of high-minority, low socio-economic status (SES) children. Low-SES children are less likely to attend preschool (Heckman, 2008; Fryer et al., 2015) and exhibit less patience and greater impulsivity than high-SES children (Watts et al., 2018). Our results assess one channel through which preschool can help narrow early skill gaps between those born poor and those born rich.

In what follows, Section 2 describes our experimental design. Section 3 characterizes the time preferences of the children. Section 4 describes the treatment effects of preschool. Section 5 discusses the relationship with later outcomes in middle childhood. Section 6 concludes.

2. Experimental Design

2.1 Chicago Heights Early Childhood Center

The children in our experiment are part of the Chicago Heights Early Childhood Center (CHECC). CHECC was a randomized study implemented in 2010-14 in Chicago Heights, Illinois that focused on investigating types of early childhood investments that can reduce the academic achievement gap. Chicago Heights is an area with 29% of people below poverty level and a median

income of $43,102. The suburb is also highly diverse, with 41.5% Black and 33.9% Hispanic or Latino households.[3]

CHECC recruited children ages 3-4 annually for 4 years, generating 4 cohorts of participants. Upon enrolling, children were randomized to preschool or to a control group for a period of 1-2 years.[4] All children participated in periodic skills assessments, including a baseline assessment and an endline assessment. The preschool programs were free, full day programs that differed on which curriculum was offered, as follows:

- Preschool: Tools of the Mind (Cohorts 1 & 2): Used the *Tools of the Mind* curriculum with a strong focus on executive functioning skills (Bodrova and Leong, 2007).
- Preschool: Literacy Express & Math (Cohorts 1 & 2): Used the *Literacy Express* curriculum that focused on building literacy skills (Lonigan, 2006; Lonigan et al., 2005), supplemented with math instruction.
- Preschool: CogX (Cohorts 3 & 4): Used a new curriculum combining elements of executive functioning and literacy, half-day instruction and half-day daycare. Included parental component where parents met with teachers bi-weekly to learn how to scaffold child's learning at home.
- Control (Cohorts 1-4): Did not receive CHECC preschool.

[3] U.S. Census Bureau Quick Facts: http://quickfacts.census.gov/qfd/states/17/1714026.html

[4] Two other programs were also part of CHECC: Parent Academy included classes for parents but did not include a child intervention (see Fryer et al., 2015), while Kinderprep included a shorter, 2-month long preschool program utilizing the CogX curriculum. Due to budget constraints, we made the difficult choice of focusing only on the Control and Preschool groups when we collected the time preference data described in this paper.

At age 5, children enrolled in kindergarten in one of the elementary schools in District 170 or in surrounding districts. We had an agreement with District 170 to continue conducting assessments during school hours with the children who ultimately enrolled in the district. As such, our experimental sample consists of all children who were in kindergarten to second grade in these schools in 2015.

There are benefits and drawbacks to assessing the children during school hours. One benefit is that it reduces the selective attrition that may arise if we asked parents to voluntarily bring children for an assessment. This type of attrition could be problematic if parents of treatment kids have stronger social ties with the CHECC program than parents of control kids, and these social ties are unobservable. However, a drawback is that since we can only assess District 170 children, we capture only 17.8% of the CHECC sample in our data. Because children typically attend school based on the neighborhood in which they live, the sample in this paper differs somewhat in terms of demographics relative to the overall sample. To account for the possibility that CHECC assignment influenced the likelihood of remaining in District 170, we use inverse probability weights (IPW) in our robustness checks.

Table 1 provides summary statistics of our sample overall and by treatment group. At the time of the experiment, children were on average 7-8 years old (SD=0.05-0.14). About half of our sample is female. Our sample is 72-80% Hispanic, 20-30% Black and 2-8% White. 67-78% of mothers had graduated high school and 23-37% worked full time. During the time preference experiment in September 2015, mothers were 34-36 years old on average. After a Sidak correction for multiple hypothesis testing, we find only one marginally significant p-value for age of child. To account for potential imbalance across groups, in our regressions we display results with and without controls for background variables. We find that results are similar across these

specifications. Appendix Table A.1 shows that the sample is balanced by income and elementary school.

One limitation of our data is the small sample size for each preschool curriculum. However, lumping all treatments into a single preschool group would limit our ability to examine the effect of curricula that emphasize self-regulation skills, compared to curricula that emphasize traditional reading and math skills. Given this trade-off, we chose to explore results by curriculum and to report Wild bootstrap-t confidence intervals to adjust our inference for the small number of treated observations (see Wu, 1965; Cameron et al., 2008; Rao, 2019).

2.2 Experimental Design

The experiment was conducted at school in September 2015 over a period of two weeks, with two decision-making days in each week (described in Table 2). We used glow-in-the dark bracelets as the reward. Glow-in-the-dark bracelets fade within 24 hours of receipt, which ensures that our reward is actually "consumed" on the day of receipt. Children were reminded about this feature of the bracelets by the experimenters on each day, so that they would not have the expectation of hoarding them across weeks. We chose glow-in-the-dark bracelets after analyzing children's preferences over these bracelets compared to other potential rewards (e.g., glow-in-the-dark sticks, temporary tattoos, stickers, figurines). We found them to be appropriate for the age group of 5-10 year olds. They were equally preferred by boys and girls.

On the first day, children chose between receiving one bracelet on the second day versus two bracelets on the third day. They recorded their decisions in private on white activity sheets that were collected by the experimenters at the end of each experiment day. As shown in Appendix B, the activity sheets featured picture representations of both the number of bracelets and time of receipt, together with written descriptions. Children were asked to circle their preferred choice

between a smaller, earlier reward and a larger, later reward. We refer to the decision made on Day 1 as the "Advance Decision", since children made their selection in advance and had to wait for whichever option they chose.

Table 1: Summary Statistics by CHECC Treatment

		Years 1-2			Years 3-4	
	All	Control	Tools	Lit	Control	CogX
child age	7.726	8.386	8.829	8.486	6.790	6.714
	(0.056)	(0.071)	(0.121)*	(0.135)	(0.051)	(0.063)
female	0.490	0.432	0.513	0.564	0.509	0.525
	(0.025)	(0.042)	(0.081)	(0.080)	(0.047)	(0.066)
black	0.240	0.215	0.179	0.231	0.268	0.288
	(0.022)	(0.035)	(0.062)	(0.068)	(0.042)	(0.059)
hispanic	0.758	0.770	0.744	0.718	0.732	0.814
	(0.022)	(0.036)	(0.071)	(0.073)	(0.042)	(0.051)
white	0.029	0.015	0.077	0.051	0.027	0.017
	(0.009)	(0.010)	(0.043)	(0.036)	(0.015)	(0.017)
mother age	34.936	35.534	33.884	34.704	34.630	34.680
	(0.513)	(0.739)	(1.535)	(1.543)	(1.180)	(1.578)
mother is HS graduate	0.710	0.696	0.720	0.682	0.781	0.667
	(0.034)	(0.052)	(0.092)	(0.102)	(0.074)	(0.114)
mother works full time	0.307	0.367	0.308	0.227	0.278	0.231
	(0.034)	(0.055)	(0.092)	(0.091)	(0.076)	(0.084)
baseline cog (std)	-0.316	-0.309	-0.545	-0.141	-0.256	-0.391
	(0.038)	(0.061)	(0.115)	(0.113)	(0.077)	(0.095)
baseline noncog (std)	-0.147	-0.023	-0.216	-0.127	-0.338	-0.103
	(0.049)	(0.077)	(0.183)	(0.152)	(0.093)	(0.128)
Observations	388	139	39	39	112	59

Notes: Standard errors in parentheses. Stars denote significance relative to control after Sidak correction for multiple hypothesis testing. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

On the second day, experimenters returned to the classroom and surprised children with the option to change their mind. That is, children who had previously preferred two rewards on the third day could instead select one reward on the second day, and vice versa. Children received back their decision sheet from the previous day so that they could review it. Then, they received a new sheet (which was yellow to separate it from the previous sheet) and were asked to circle their preference. We operationalized the re-allocation in this way in order to assure that equal

effort was required to remain with the previous choice versus to switch to a new choice, and in order to assure privacy (every child received and returned a yellow sheet, regardless of decision made). We refer to the decision on Day 2 as the "Immediate Decision", since the one bracelet option would be received immediately if selected.

Week 2 differs from Week 1 in one crucial way. After making the Advance Decision, participants had the opportunity to commit to this decision, or to keep open the possibility of changing their mind in the Immediate Decision the next day. We operationalized the commitment choice by asking students to either state whether they would like their current decision sheet to be the final decision, or whether they would like to make their final decision the next day on their yellow sheet. To decide, children either circled a white or a yellow box on their activity sheets. We presented the commitment device after the Advance Decision so that the commitment device would only affect inconsistency and impulsivity in Week 2, and not patience.

This design allows us to distinguish between specific components of time preferences. From Week 1, we identify patience from the Advance Decision, when both options were paid in the future and present-bias did not influence selections. We identify inconsistency based on whether children change their mind in going from the Advance Decision in Day 1 to the Immediate Decision in Day 2. Impulsivity is a subset of inconsistency, occurring when children initially chose the later, two-bracelet option in the Advance Decision but switched to the sooner, one-bracelet option when it was paid out immediately in the Immediate Decision. In Week 2, participants made the same allocation decision in Advance and Immediate Decisions, allowing us to assess how patience, inconsistency, and impulsivity changed from Week 1 to Week 2.

We classify children who make the same selection in Advance Decisions as "stable" and children who make the opposite selection in Advance Decisions as "unstable". This "stability of

patience", which has been under-explored in prior work on time preferences, is a key feature of the analysis in Sections 3-5. It is especially relevant for examining how time preferences can help children invest in later human capital, since it captures sustained endurance at delaying gratification over long periods of time. Such behavior requires multiple components. First, it requires setting a goal – in the case of our experiment, choosing the later option in the Advance Decisions. Second, it requires being time consistent, following through, and holding out for the later option in the face of temptation. This is measured by the transition between the Advance to the Immediate Decisions within each week. It is also measured by the Advance Decision in Week 2, when children know the task structure and must decide whether to commit to holding out again for the later option.

We always conducted the experiment in the last 15-30 minutes of the school day. This design feature assured that children would be able to immediately receive their rewards after participation (rather than, for example, participating in the morning and having to wait until the end of school to get their rewards). Experimenters were blind to whether participating children had been in the Preschool or Control groups, and children from both groups participated in each session. Experimenters read the instructions out loud to the group, and students also completed a quiz on understanding during the first day of the study (see practice book in Appendix B).

At the end of the experiment on the second day, all children received an envelope, which either contained one glow-in-the-dark bracelet (for students who selected the earlier, smaller reward) or was empty (for students who selected the later, larger reward). On the third day, all participants again received an envelope, which either contained two glow-in-the-dark bracelets (for students who selected the later, larger reward) or was empty (for students who selected the earlier, smaller reward). We gave all participants envelopes in each payout period to ensure

privacy of decisions. Despite our efforts, there are many ways in which children could have shared information regarding their decisions. Children were free to wear the bracelets afterwards, and could have been observed doing this by their classmates. They could have also discussed their choices with their classmates throughout the experiment. We will present evidence in Section 4 that children appear to learn of their classmates' behavior in Week 1 and to be influenced by them in Week 2.

2.2 Experimental Procedures

Table 1: Experiment Design and Procedures

Week 1		
Day 1	**Day 2**	**Day 3**
* Choose between: 1 reward on Day 2 or 2 rewards on Day 3	* Can make changes to Day 1 decision * Receive reward if chosen for Day 2	* Receive reward if chosen for Day 3
Week 2		
Day 1	**Day 2**	**Day 3**
* Choose between: 1 reward on Day 2 or 2 rewards on Day 3 * Decide whether to commit	* Can make changes to Day 1 decision, if didn't commit * Receive reward if chosen for Day 2	* Receive reward if chosen for Day 3

3. Characterizing Time Preferences

In this Section, we characterize time preferences for the full sample of 388 children. Appendix Table A.2 reports the raw numbers and percentages. We first focus on Advance Decisions, which we use to measure patience. Among the full sample of 388, 81% and 76% chose the delayed option of two bracelets in Weeks 1 and 2, respectively. Using a paired t-test, we find that children are significantly less likely to choose the later option in Week 2 than Week 1 ($p<0.0001$), displaying

less patience over time. We next dig into this unanticipated result by investigating the stability of patient choices across weeks.

We classify children by whether they were stable or unstable in patience by comparing their Advance Decisions between Week 1 and Week 2. 56% (216) were *stable in patience*, choosing the later option of two bracelets in both weeks. 7% (28) were *stable in impatience* and chose the sooner option of one bracelet in both weeks. 24% (92) exhibited instability in time preferences. 14% (54) exhibit *instability in patience* and switch from the later option to the sooner option from Week 1 to Week 2. Lastly, 10% (38) exhibit *instability in impatience* and switched from the sooner to the later option from Week 1 to Week 2. The remaining children were not present for both questions, so we could not assess the stability of their preferences across weeks.

We next examine the consistency of choices within weeks, where children first make the Advance Decision but can change their mind in the Immediate Decision the next day. In the full sample, 21% (77) of children exhibit inconsistency in Week 1, where their Advance and Immediate Choices differ. Of these children, 74% (57) are impulsive, switching from two bracelets in the Advance Decision to one bracelet in the Immediate Decision. By comparison, in Week 2 18% (29) of children are inconsistent, with only 55% of them (16) behaving impulsively. Using a paired t-test, we find that children do not exhibit a statistically significant difference in inconsistency across weeks ($p>0.10$), but are significantly less likely to behave impulsively ($p<0.05$) in Week 2 compared to Week 1.

Is the lower rate of impulsivity in Week 2 driven by the use of commitment devices? About half (51%) of all children use the commitment device. Of the 244 children who successfully end up with the later option in Week 2, 55% (135) used the commitment device. Given that only half

of these "successful" children used the commitment device, the commitment device does not appear crucial to obtaining two bracelets at the end of Week 2.

4. Treatment Effects

We next demonstrate that our Preschool intervention improved patient behavior among children. However, this did not occur because Preschool automatically made children more patient, since there was no difference in behavior between Preschool and Control children when they initially participated in the glow bracelet task. Rather, we observe greater patience in the Preschool group in Week 2. The effects are driven by the *Tools* curriculum, which focused on self-regulation skills. Results are similar between the Control group and children who were randomized to the *Literacy* curriculum, which placed greater emphasis on traditional literacy and math skills. For children in the *CogX* curriculum, which was a hybrid version of *Tools* and *Literacy*, estimates tend to fall between estimates for the Control and *Tools* groups. However, as with the *Literacy* group, results for the *CogX* group do not tend to significantly differ from the Control group.

We begin with the striking finding that patience declined over time, with fewer children selecting the later option in Week 2 than Week 1 in the overall sample ($p<0.0001$). We explore how these patterns differ by CHECC treatment using the following specification:

$$choice_{it} = \alpha_0 + \alpha_1 1(week_t = 2) + \alpha_2 D_i + \alpha_3 D_i 1(week_t = 2) + \alpha_4 X_i + \epsilon_{it} \qquad (1)$$

where $choice_{it}$ is child $i's$ choice in week t, D_i is a matrix of dummies for Preschool curriculum, and X_i represents a matrix of child-specific controls including age during the glow bracelet task, year(s) of CHECC participation, gender, race, and baseline cognitive and noncognitive scores prior to randomization into CHECC treatment. The matrix X_i also may include parental controls, namely the mother's age, high school graduation status, and full time employment status. The idiosyncratic error component ϵ_{it} varies by both child i and week t. In this specification, as in almost all

specifications that follow, we cluster standard errors at the child level since CHECC treatment was randomized at the child level (Abadie et al., 2022). As mentioned above, we report 90% confidence intervals after using the Wild bootstrap-t procedure to account for the small number of treated children per preschool curriculum.

Table 3 reports the regression results on choices between Week 1 and Week 2. Panel A shows estimates for the likelihood of choosing the later option in the Advance Decision, while panel B shows estimates for the likelihood of choosing the later option in the Immediate Decision. We find no differences across groups in the likelihood of making the patient choice in Week 1, as the coefficients on the *Tools, Literacy,* and *CogX* curricula are all small and insignificant. However, the coefficient on Week 2 is consistently negative and significant ($p<0.05$), indicating that children are 9-10 percentage points less likely to make the patient choice in Week 2 compared to Week 1. The exception to this pattern is the *Tools* group. Those in the *Tools* group are 16 percentage points more likely to choose the delayed option in Week 2 compared to the Control group (adding the coefficients on $Tools$ and $Tools \times Week\ 2$, $p<0.01$).

Panel B then examines the Immediate Decision asked on Day 2 of each week, where children could receive the sooner option immediately but still had to wait if they chose the later option. For Week 2, we focus on the likelihood of *ending up* with the later option, through either using the commitment device in the Week 2 Advance Decision or not committing and choosing the later option in the Week 2 Immediate Decision. As with the Advance Decision, we find little evidence of differences by group in Week 1. Although the likelihood of choosing the later option still grows from Week 1 to Week 2 in the *Tools* group (by 14 percentage points, $p<0.10$), it no longer significantly declines for the Control group. The net result is that, in contrast to the Advance Decision, the *Tools* group no longer outperforms the Control group in Week 2 for the Immediate

Decision (adding the coefficients on $Tools$ and $Tools \times Week\ 2$, p>0.10). Appendix Table A.3 presents results on choosing the later option for only children who did not commit in the Week 2 Advance Decision. The results are similar whether or not we exclude children who commit, suggesting that the results for the Immediate Decision are not driven by use of the commitment device for *Tools* children.

4.1 Patience and Stability

To further investigate why the *Tools* curriculum appeared to improve patience in Week 2 but not Week 1, we plot the predicted rates of patient choices in the Advance Decisions after controlling for age, year of CHECC participation, gender, race, baseline cognitive and noncognitive scores, and parental controls. As with Table 3, we adjust the 90% confidence intervals using wild bootstrapping. The two leftmost bars on Figure 1 show the proportion of patient choices in Week 1 and Week 2. The predicted proportion who choose the later option significantly decline in the Control group (p<0.01), but directionally rise in the *Tools* group.

It thus appears that the gap between *Tools* and Control emerges in Week 2 but not in Week 1 because the *Tools* and the Control group move in opposite directions. To investigate why, we again classify children into four categories based on how they switch between Week 1 and Week 2 in their Advance Decisions. Those who are *stable in patience* choose the later option in both weeks; those who are *stable in impatience* choose the sooner option in both weeks. Children who are *unstable in patience* choose the later option in Week 1 and switch to the sooner option in Week 2. Finally, children who are *unstable in impatience* choose the sooner option in Week 1 and switch to the later option in Week 2.

The next four bars in Figure 1 show the predicted proportions by CHECC treatment. Between the Control and *Tools* groups, we find no significant difference in the proportion of subjects who are stable in patience. However, those in the Control group are 7% more likely to be stable in impatience ($p<0.10$). Subjects who are stable cannot account for the change in

Table 3: Regression of Advance and Immediate Decisions on CHECC Treatment

	Chose Later Option			
	(1)	(2)	(3)	(4)
	A: Advance Decision			
Week 2	-0.092	-0.094	-0.095	-0.096
	[-0.149,-0.035]**	[-0.151,-0.037]**	[-0.151,-0.039]***	[-0.152,-0.040]**
Tools	-0.014	-0.024	-0.020	-0.000
	[-0.128,0.100]	[-0.139,0.090]	[-0.137,0.098]	[-0.111,0.110]
Tools × Week 2	0.180	0.182	0.183	0.184
	[0.056,0.306]**	[0.058,0.308]**	[0.059,0.309]**	[0.059,0.309]**
Literacy	-0.015	-0.028	-0.032	0.010
	[-0.126,0.097]	[-0.141,0.084]	[-0.148,0.083]	[-0.104,0.124]
Literacy × Week 2	0.014	0.014	0.013	0.014
	[-0.136,0.168]	[-0.136,0.167]	[-0.136,0.165]	[-0.137,0.168]
CogX	-0.038	-0.043	-0.037	-0.041
	[-0.153,0.078]	[-0.162,0.076]	[-0.160,0.085]	[-0.159,0.080]
CogX × Week 2	0.149	0.153	0.153	0.152
	[0.008,0.291]*	[0.009,0.295]*	[0.010,0.296]*	[0.009,0.292]*
Observations	707	707	707	707
	B: Immediate Decision, inc. Committers			
Week 2	0.002	0.001	0.002	0.003
	[-0.061,0.067]	[-0.062,0.066]	[-0.061,0.067]	[-0.060,0.067]
Tools	-0.030	-0.038	-0.044	-0.010
	[-0.159,0.101]	[-0.166,0.092]	[-0.176,0.088]	[-0.126,0.109]
Tools × Week 2	0.145	0.145	0.146	0.142
	[0.021,0.268]*	[0.021,0.267]*	[0.021,0.268]*	[0.020,0.264]*
Literacy	-0.097	-0.109	-0.130	-0.068
	[-0.237,0.041]	[-0.254,0.033]	[-0.270,0.014]	[-0.193,0.061]
Literacy × Week 2	0.042	0.044	0.045	0.052
	[-0.109,0.193]	[-0.107,0.194]	[-0.106,0.195]	[-0.102,0.206]
CogX	-0.014	-0.002	0.021	0.023
	[-0.145,0.117]	[-0.134,0.130]	[-0.116,0.154]	[-0.095,0.143]
CogX × Week 2	0.119	0.115	0.118	0.121
	[-0.049,0.284]	[-0.054,0.286]	[-0.050,0.287]	[-0.044,0.289]
Observations	689	689	689	689
Child Gender & Race		Yes	Yes	Yes
Pre-assessment scores			Yes	Yes
Parental Controls				Yes

Notes: All regressions control for year of CHECC participation and child age. Parental controls include age, HS graduate status, and full-time employment of mother and father. Dummies for missing variables suppressed. Standard errors clustered by child. 90% confidence intervals use wild bootstrap-t method from Cameron, Gelbach, and Miller (2008). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

patient choices from Week 1 to Week 2, so we next move to subjects who are unstable. We find that children in the Control group are more likely to be unstable in patience than the *Tools* group. In particular, those in the Control group are 11% more likely than the *Tools* group to choose the later option in Week 1 but switch to the sooner option in Week 2 ($p<0.10$).

The *Literacy* and *CogX* groups do not significantly differ from the Control group on the various patience measures shown in Figure 1. For the *Literacy* group, these results are unsurprising given that the focus of the curriculum was on math and reading, which may not necessarily translate to greater patience. Point estimates for the *CogX* group tend to fall between the *Literacy* and *Tools* groups, which makes sense given that *CogX* was a hybrid of the *Literacy* and *Tools* curricula. However, we refrain from drawing conclusions regarding the effects of the *CogX* curriculum, since we do not have sufficient evidence to establish a significant gap between *CogX* and Control children.

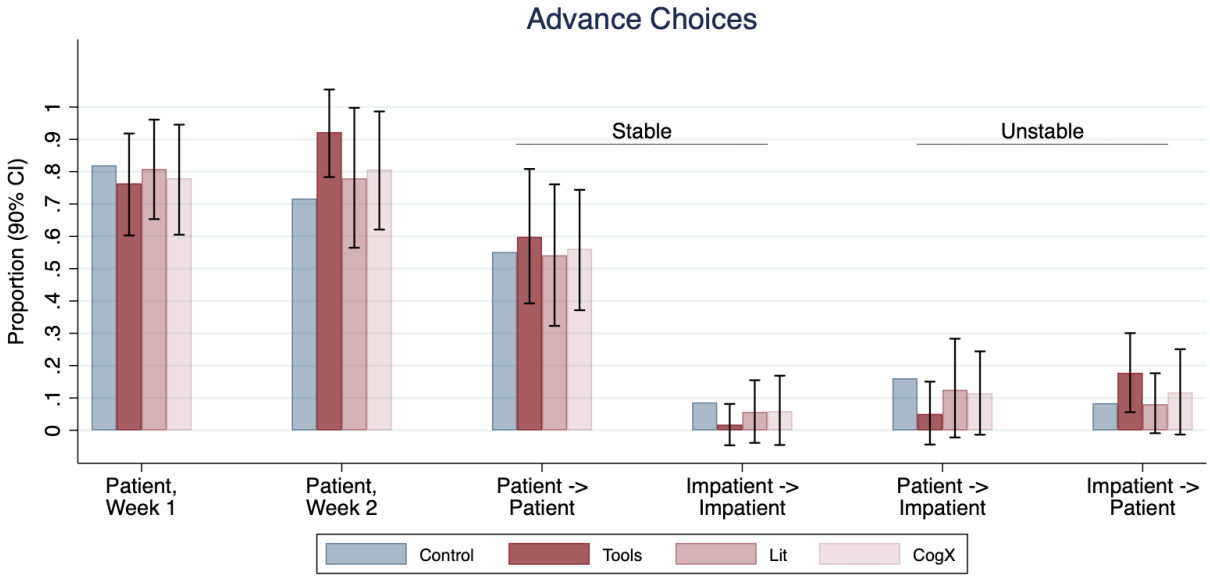


Figure 1: Stability of patience across weeks

4.2 Time Inconsistency

Why might children in the Control group change from the delayed option in Week 1 to the sooner option in Week 2, demonstrating greater impatience over time? We explore choices within Week 1, particularly on whether children change their mind in moving from the Advance Decision to the Immediate Decision. It is possible that children who initially chose the later option in the Advance Decision changed their mind in the Immediate Decision, where the sooner option would be paid out immediately. This impulsive behavior in Week 1 could then influence their Advance Decision in Week 2. For example, impulsivity in Week 1 may lower children's beliefs about their ability to wait for the later option, leading them to choose the earlier option in Week 2. Alternatively, receiving the one bracelet in Week 1 may lead children to convince themselves that they prefer the earlier option, due to endowment or cognitive dissonance effects. Differences in Week 1 behavior could thus account for the gap in Week 2 patience between the Control and *Tools* groups.

We test this hypothesis by examining inconsistency and impulsivity within week, as shown in Figure 2. We find no differences in inconsistency, impulsivity, or commitment behavior across CHECC treatments in either week, indicating that the difference in Week 2 patience between the *Tools* and Control groups are unlikely driven by impulsivity in Week 1.[5] We next examine the specific group of children who exhibit instability in patience, since our prior results suggest that it is this group that drove the gap in Week 2 patience between the *Tools* and Control groups. Of the children in the Control group who exhibit instability in patience, very few (27%) behaved impulsively in Week 1. The results indicate that inconsistent behavior in Week 1 cannot explain the Control group's instability in patience, since the majority of Control children (73%)

[5] For Week 2, we report estimated inconsistency and impulsivity among children who did not use the commitment device.

consistently chose the later option in both the Advance and Immediate Decisions in Week 1 before switching to the sooner option in Week 2.

Furthermore, we find no significant relationship between inconsistent behavior in Week 1 and patience in Week 2. As shown in Appendix Table A.4, our estimated correlations are close to 0, both on the overall sample as well as for each CHECC treatment. This null result provides further evidence that time inconsistency is not driving the drop in patience in Week 2 among the Control group.

4.3 Social Environment

As noted above, the majority of children who were unstable in patience actually chose the later option in *both* the Advance and Immediate Decisions of Week 1. This was surprising, since we would expect children who successfully waited for the larger reward in Week 1 to continue to exhibit patience in Week 2. Given that own behavior poorly predicts patience in Week 2, we turn to external explanations. We exploit quasirandom variation in the assignment of children to elementary classrooms to determine whether social environment could have played a role.

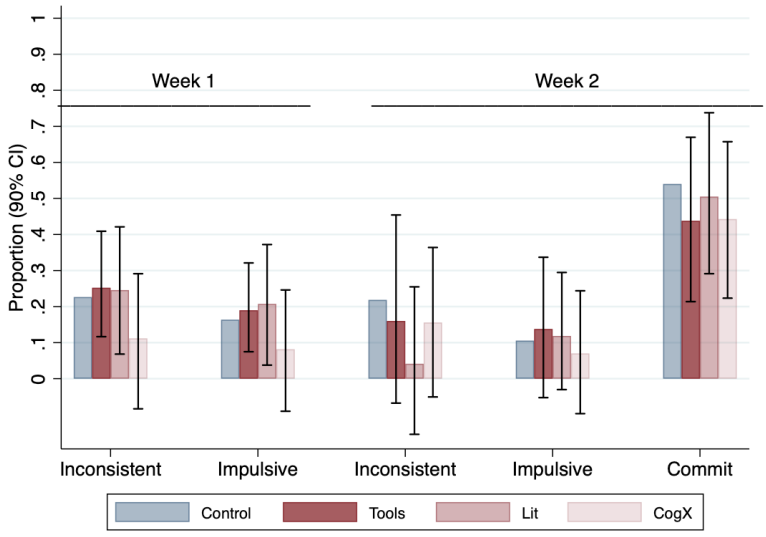


Figure 2: Inconsistency and impulsivity within weeks

Following participation in CHECC, children enter elementary school in the Chicago Heights School District. Differences across classrooms should be orthogonal to their CHECC treatment status, since teachers assign them to classrooms without knowing their prior Preschool curriculum assignment. In the balance tests shown in Appendix Table A.1, we find no significant differences between any preschool curriculum and Control in the likelihood of enrolling in any of the 7 schools in the Chicago Heights School District. In our decomposition exercise below, we also verify that our results are not driven by significant differences in baseline characteristics between the *Tools* and Control groups.

We investigate whether children's Week 2 decisions depend on the share of their classmates who ended up with the later option of two bracelets. Figure 3 plots the linear prediction of Week 2 patience based on the share of classmates with two bracelets in Week 1. Panel A focuses on comparing the Control and *Tools* groups. It shows that Control children were more likely to behave impatiently in Week 2 if they happened to be in classes where more of their classmates ended up with the sooner option in Week 1. Out of the explanations we consider, this is the only one that explains why many of the Control group chose the sooner option in Week 2 despite choosing the later option in both days of Week 1.

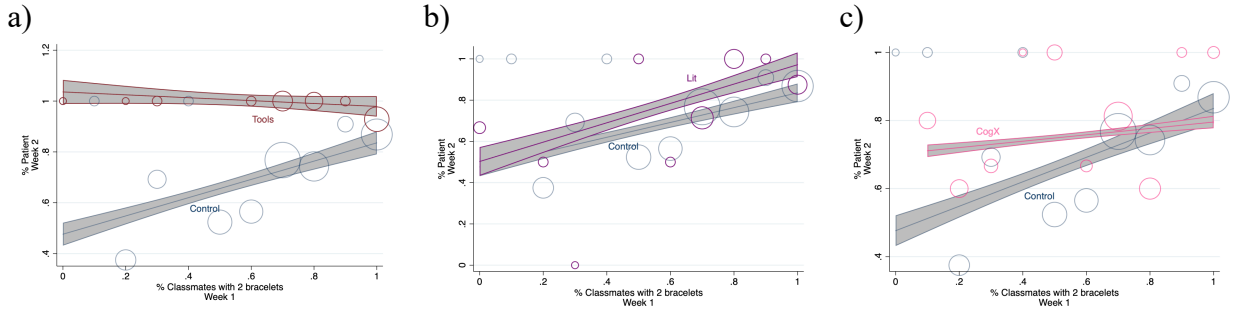


Figure 3: Patience in week 2 and proportion of patient classmates in week 1. Panel a presents 95% CI of linear predictions for Control and Tools. Panel b presents linear predictions for Control and Literacy. Panel c presents linear predictions for Control and CogX.

In contrast, children from the *Tools* group almost always choose the later option in Week 2, no matter their classmates' outcomes in Week 1. Figure 3 suggests that differences in susceptibility to peer influence could contribute to why patience significantly drops among the Control group from Week 1 to Week 2, leading to significantly greater patience in the *Tools* group in Week 2 but not Week 1. Panels B and C estimate Week 2 patience for the *Literacy* and *CogX* curricula, respectively. They show that both the *Literacy* and *Cogx* curricula appear to be influenced by the share of classmates with two bracelets. Like with the other measures above, the *Literacy* group resembles the Control group. The CogX group is statistically indistinguishable from the Control group, but the estimated susceptibility to peer influence falls between that of the *Tools* and Control groups.

Motivated by the evidence from Figure 3, we perform a Oaxaca-Blinder decomposition to determine how much of the gap in Week 2 patience between the *Tools* and Control groups can be attributed to differences in susceptibility to social environment. More precisely, the decomposition distinguishes between differences in "coefficients" – responsiveness to peer influence – and differences in "endowments" – in this case, whether the *Tools* and Control groups see different shares of classmates with two bracelets. Second, the decomposition allows us to compare the relative importance of peer influence with other potential predictors of patient behavior.

We specify the decomposition in the following way:

$$y^T - y^C = \alpha_0^T - \alpha_0^C + \underbrace{(\alpha_1^T - \alpha_1^C)X^C}_{coefficients} + \underbrace{\alpha_1^C(X^T - X^C)}_{endowments} + \underbrace{(\alpha_1^T - \alpha_1^C)(X^T - X^C)}_{interaction} + \epsilon^T - \epsilon^C \quad (2)$$

where y^g for $g \in \{T, C\}$ is the proportion who choose the later option in Week 2. The matrix of predictors X^g consists of the share of classmates with two bracelets in Week 1, age, race, gender, cognitive ability, non-cognitive ability, and parental characteristics. The term $(\alpha_1^C - \alpha_1^C)X^C$

represents differences in "coefficients". It reflects the gap that is attributable to differences in how the *Tools* and Control group respond to predictors. The term $\alpha_1^C(X^T - X^C)$ represents the difference in "endowments", namely whether the predictors in the *Tools* and Control groups systematically differ. Lastly, $(\alpha_1^T - \alpha_1^C)(X^T - X^C)$ captures interactions between the "endowments" and "coefficients" terms (Fairlie, 2005; Jann, 2008).

The main independent variable of interest is the share of a child's classmates with two bracelets out of all classmates that participated in the delay of gratification task. Given our assumption that the fraction of classmates with two bracelets is as-good-as-random, the unit of treatment assignment in this case is at the classroom level. We therefore cluster at the classroom level, rather than the child level as in our other analyses (Abadie et al., 2022).

Table 4 presents the results of the Oaxaca decomposition. The first column presents overall breakdowns: the predicted shares of children who exhibit patience in Week 2 is 75.7% among the Control group and 97% among the Tools group, yielding a difference of 21.3 percentage points. Of these 21.3 percentage points, 21.2 percentage points (99.5%) arise from differences in responsiveness to predictors between the Control and *Tools* groups.

Columns (2), (3), and (4) further demonstrate the importance of peer behavior compared to other potential explanations. Column (3) estimates the role of each predictor in contributing to the "coefficients" term. We find that the Control group's greater susceptibility to peer influence plays a large role in explaining the overall 21.3 percentage point difference in Week 2 patience between the Control and *Tools* groups. In fact, susceptibility to peer influence appears to be the *only* significant predictor of the gap in Week 2 patience. Coefficient estimates on all other potential predictors – children's demographic characteristics, maternal characteristics, and baseline assessment scores – are small and statistically insignificant. The exception is that the coefficient

on female is marginally significant ($p<0.10$), suggesting that the *Tools* group might have generated gender differences in patience.

The interpretation is that children in the Control group would exhibit significantly greater patience in Week 2 if they were as independent in the face of peer influence as those in the *Tools* group. Equivalently, if the *Tools* group was as susceptible to peer influence as the Control group, each 1 percentage point decline in the share of patient classmates would lower the child's likelihood of exhibiting patience in Week 2 by 0.7 percentage points ($p<0.001$). The Oaxaca decomposition indicates a large difference in responsiveness to peer influence: Control children exhibit substantial elasticity to the patience of their classmates, but the estimated elasticity among *Tools* children is small and indistinguishable from 0. These results are consistent with research in child development showing that the *Tools of the Mind* curriculum reduced bullying and peer ostracism (Diamond et al., 2019), which may free children from pressure to conform.

We run similar Oaxaca decompositions for the *Literacy* and *CogX* groups. Appendix Tables A.5 and A.6 show that the gap in Week 2 patience is 7.1 and 8.1 percentage points for the *Literacy* and *CogX* groups, which are only a third of the gap for *Tools* children and statistically insignificant. Given that patience does not significantly differ between *Literacy, CogX*, and Control children, it is not surprising that responses to peer influence also do not significantly differ.

Our results suggest that the *Tools* curriculum fosters greater patience over time by making children more independent of their social environment. Despite what their classmates choose in Week 1, children from the *Tools* group are more likely to persist in waiting for the larger reward, the option that was initially preferred by the majority of children in Week 1. On the other hand, children in the Control group are more likely to choose the earlier, smaller reward if more of their classmates received it in Week 1. This is the case even for children who successfully waited for

the later reward on both days of Week 1. We view these findings as evidence that social environment plays an important role in shaping patient behavior, just as it does many other behaviors. Interventions which teach waiting strategies absent a social environment that accounts for interactions between children may miss out on key determinants of how patient behaviors form over time. Our findings support the recommendation in Falk et al. (2020) that interventions to improve delay of gratification center around the social environment in which preferences form, rather than focusing narrowly on waiting strategies in a vacuum.

Table 4: Oaxaca decomposition on % classmates with 2 bracelets

	Choose Later Option, Week 2 Advance Choice Tools vs. Control			
	Overall			
	(1)	(2)	(3)	(4)
Control	0.757***			
	(0.0563)			
Tools	0.970***			
	(0.0292)			
Control - Tools	-0.213***			
	(0.0633)			
Endowments	0.0330			
	(0.0440)			
Coefficients	-0.212**			
	(0.0866)			
Interaction	-0.0339			
	(0.0858)			
Decomposition:		Endowments	Coefficient	Interaction
% classmates with 2 bracelets		0.00876	0.742***	-0.0542
		(0.0108)	(0.198)	(0.0540)
child age		-0.00896	0.742	-0.0344
		(0.0305)	(1.130)	(0.0547)
female		0.0119	0.120*	-0.0175
		(0.0205)	(0.0637)	(0.0292)
black		0	0.0108	0.00150
		(.)	(0.0528)	(0.00913)
hispanic		0.00137	0.120	-0.00155
		(0.0125)	(0.222)	(0.0145)
baseline cog (std)		0.0217	0.0705	-0.0252
		(0.0275)	(0.0585)	(0.0326)
baseline noncog (std)		-0.0173	-0.00934	0.0117
		(0.0185)	(0.0166)	(0.0206)
mother's age		0.0115	0.0259	-0.00553
		(0.0235)	(0.270)	(0.0578)
mother is HS grad		0.0180	0.0402	-0.00809
		(0.0236)	(0.0997)	(0.0220)
mother works full-time		-0.000582	0.0132	-0.000307
		(0.00811)	(0.0254)	(0.00431)
N		144		

Standard errors clustered by classroom in parentheses. All regressions additionally control for years of CHECC participation and dummies for missing variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Regression of Impulsivity on CHECC Treatment

	Impulsivity			
	(1)	(2)	(3)	(4)
Week 2	-0.034	-0.034	-0.035	-0.036
	[-0.109,0.042]	[-0.109,0.042]	[-0.109,0.042]	[-0.110,0.041]
Commit	0.005	0.001	0.004	0.013
	[-0.074,0.082]	[-0.077,0.080]	[-0.075,0.084]	[-0.065,0.092]
Week 2 × Commit	-0.113	-0.112	-0.111	-0.111
	[-0.204,-0.020]**	[-0.204,-0.019]**	[-0.203,-0.019]**	[-0.203,-0.018]**
Tools	-0.082	-0.090	-0.086	-0.085
	[-0.141,-0.024]**	[-0.150,-0.031]**	[-0.150,-0.022]**	[-0.155,-0.016]*
Tools × Week 2	0.105	0.105	0.104	0.106
	[-0.119,0.347]	[-0.103,0.353]	[-0.087,0.363]	[-0.091,0.327]
Tools × Week 2 × Commit	-0.116	-0.116	-0.115	-0.117
	[-0.078,0.208]	[-0.088,0.189]	[-0.103,0.187]	[-0.095,0.181]
Literacy	0.118	0.128	0.145	0.125
	[-0.022,0.248]	[-0.022,0.247]	[-0.023,0.247]	[-0.021,0.248]
Literacy × Week 2	-0.190	-0.190	-0.189	-0.185
	[-0.429,0.049]	[-0.428,0.050]	[-0.428,0.050]	[-0.423,0.049]
Literacy × Week 2 × Commit	0.225	0.226	0.224	0.210
	[-0.273,0.035]	[-0.272,0.043]	[-0.268,0.041]	[-0.269,0.041]
CogX	0.067	0.054	0.045	0.045
	[-0.322,0.094]	[-0.321,0.095]	[-0.320,0.094]	[-0.322,0.092]
CogX × Week 2	-0.119	-0.113	-0.112	-0.112
	[-0.031,0.484]	[-0.032,0.486]	[-0.031,0.482]	[-0.040,0.466]
CogX × Week 2 × Commit	0.160	0.154	0.154	0.153
	[-0.049,0.368]	[-0.058,0.365]	[-0.058,0.364]	[-0.059,0.364]
Observations	663	663	663	663
Child Gender & Race		Yes	Yes	Yes
Pre-assessment scores			Yes	Yes
Parental Controls				Yes

Notes: All regressions control for year of CHECC participation and child age. Parental controls include age, HS graduate status, and full-time employment of mother and father. Dummies for missing variables suppressed. Coefficient on commitment in Week 1 suppressed (insignificant). Coefficients on treatment and commitment interaction suppressed (all insignificant). Standard errors clustered by child. 90% confidence intervals use wild bootstrap-t method from Cameron, Gelbach, and Miller (2008). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.4 Commitment Demand

Finally, we investigate take-up of the commitment device, which prevents children from changing their minds between the Advance and Immediate Decisions in Week 2. We first investigate whether the commitment device decreases subsequent impulsive behavior in Table 5. The coefficient on the interaction between Week 2 and commitment indicates that children who choose the commitment device are 11% less likely to behave impulsively ($p<0.05$). The decline in impulsivity from Week 1 to Week 2 is significant in the Control, *Tools,* and *CogX* groups ($p<0.01$

for all three groups), but not the *Literacy* group. We also test for differences in the rate of decline and find no significant differences between the Control group and any of the treatment groups. All estimates hold after the inclusion of demographic variables, baseline cognitive scores, baseline noncognitive scores, and parental controls.

Even if commitment decreases impulsivity *ex post*, children may not be aware of the value of the commitment device *ex ante*. It is therefore another matter whether children understand the commitment device and use it strategically. Table 6 reports on the relationship between behavior in Week 1 and commitment in Week 2. Despite the low rates of impulsivity among the *Tools* group, those who do behave impulsively in Week 1 are 41-50% more likely to use the commitment device in Week 2 ($p<0.10$). Appendix Table A.7 decomposes this behavior into choices made in Week 1. It shows that *Tools* children who chose the sooner option in the Week 1 Immediate Decision are 35-41% more likely to use the commitment device ($p<0.05$), suggesting that commitment was demanded by *Tools* children who had trouble delaying in the Immediate Decision. It appears that commitment was higher among the very children who were likely to benefit from it: those who failed to wait for the later option despite preferring it initially. These results are consistent with prior work documenting greater use of commitment among dynamically inconsistent adults (Ariely and Wertenbroch, 2002; Ashraf et al., 2006; Kaur et al., 2010; Augenblick et al., 2015), but contrast with other studies who find greater commitment demand among those who were dynamically consistent, and the least likely to need it (Sadoff et al., 2020; Alan and Ertac, 2015).

Remarkably, the *Tools* group is the *only* group which exhibits this behavior. We find no correlation between Week 1 behavior and commitment use among Control, *Literacy*, or *CogX*

children. Taken together, we find little evidence of strategic use of the commitment device for groups other than the *Tools* group.

Table 6: Regression of commitment on impulsive behavior in Week 1

	Commitment in Week 2			
	(1)	(2)	(3)	(4)
Impulsivity in week 1	0.047	0.013	0.022	0.031
	[-0.127,0.220]	[-0.166,0.189]	[-0.154,0.197]	[-0.151,0.212]
PK Tools	0.047	0.013	0.022	0.031
	[-0.318,0.037]	[-0.322,0.038]	[-0.334,0.026]	[-0.321,0.056]
PK Literacy	0.047	0.013	0.022	0.031
	[-0.189,0.159]	[-0.181,0.167]	[-0.203,0.154]	[-0.195,0.175]
PK CogX	0.047	0.013	0.022	0.031
	[-0.207,0.114]	[-0.217,0.103]	[-0.201,0.134]	[-0.207,0.131]
Impulsive in week 1 × PK Tools	0.413	0.502	0.492	0.520
	[0.101,0.639]*	[0.161,0.693]*	[0.153,0.683]*	[0.164,0.781]*
Impulsive in week 1 × PK Lit	-0.167	-0.118	-0.109	-0.090
	[-0.653,0.386]	[-0.617,0.461]	[-0.626,0.472]	[-0.612,0.514]
Impulsive in week 1 × PK CogX	-0.291	-0.212	-0.221	-0.193
	[-0.618,0.035]	[-0.537,0.113]	[-0.543,0.104]	[-0.541,0.147]
Observations	335	335	335	335
Child Gender & Race		Yes	Yes	Yes
Pre-assessment scores			Yes	Yes
Parental Controls				Yes

Notes: All regressions control for year of CHECC participation and child age. Parental controls include age, high school graduate status, and full-time employment of mother and father. Dummies for missing variables suppressed. Standard errors clustered by child. 90% confidence intervals use wild bootstrap-t method from Cameron, Gelbach, and Miller (2008). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.5 Robustness

First, we incorporate inverse probability weights to account for the possibility that randomized assignment to the Preschool and Control groups influenced the likelihood of remaining in the Chicago Heights School District, and therefore participating in the glow bracelet task. Second, not all children were present for all decisions in the glow bracelet task. It is possible that our results across weeks could be driven by differences in the composition of children across different days. We show our results replicate when using only the sample of children that were present for all decisions. In all regression tables, we report conventional standard errors, clustered by child, in parentheses. Beneath them, we report 90% confidence intervals after wild bootstrapping in brackets. Relative to the wild bootstrap procedure, conventional standard errors

tend to overreject the null hypothesis of no effect. Where results differ, we prefer the more conservative inference method of wild bootstrapping, given the small sample within each treatment cell in our experiment.

Appendix Tables A.8-A.9 reproduce Table 5, which explores differences across weeks in the Advance and Immediate Decisions. Across all specifications, we replicate the main results of Table 5. First, Appendix Table A.8 shows that with respect to Advance Decisions there is a significant decline in patience from Week 1 to Week 2 ($p<0.05$). The exception is the *Tools* group, which is significantly more patient than the Control group in Week 2 ($p<0.05$). Second, Appendix Table A.9 shows that with respect to Immediate Decisions we again find significantly greater rates of delaying in Week 2 in the *Tools* group compared to the Control group ($p<0.10$). For all results, point estimates grow substantially when inverse probability weights are incorporated but do not significantly change when we restrict the sample to children who were present for all decisions.

Appendix Figure A.1 reports the analysis displayed in Figure 1 after incorporating inverse probability weights (panel A) or restricting the sample to children present for all decisions (panel B). Rather than the predicted likelihoods summarized in Figure 1, we report the estimated difference relative to the Control group. For both specifications, the Control group continues to exhibit a significant 9 percentage point drop in patience from Week 1 to Week 2 ($p<0.05$). This leads to significantly greater patience among *Tools* children than Control children by Week 2 (18-22 percentage points, $p<0.01$). After these sample adjustments, we find that Control children continue to exhibit greater stability of impatience, choosing the sooner option in both weeks, than *Tools* children (5-7 percentage points, $p<0.05$). Moreover, Control children continue to be more unstable in patience, switching from the later option in Week 1 to the sooner option in Week 2, relative to *Tools* children (13 percentage points, $p<0.05$).

Finally, we reproduce our key commitment result after making these adjustments. Appendix Table A.10 shows that only *Tools* children increase commitment take-up after behaving impulsively ($p<0.10$). Commitment use continues to be uncorrelated with past behavior for Control, *Literacy*, and *CogX* children. Coefficient estimates are similar after the incorporation of inverse probability weights and after restricting the sample to children present for all decisions.

5. Relationship with Later Outcomes

In the prior section, we present evidence that preschool – in particular, the *Tools of the Mind* curriculum – improved delay of gratification among children in later years. We explore the importance of this result by linking behavior in the delay of gratification task to future outcomes collected each year: state math examination scores in 2015-2019, state reading examination scores in 2015-2019, working memory assessments in 2016-2019, attention assessments in 2016-2019, and teacher-reported disciplinary behavior in 2015-2019. In our preferred specification, we use exploratory factor analysis to extract summary measures for each set of outcomes. We then regress these factors on delay of gratification behavior, CHECC preschool treatment, and the interaction. Table 7 reports the results. For brevity, we suppress coefficients on interactions, treatment dummies, and dummies for missing variables.

We find that children who can delay in the Immediate Decision at ages 5-10 perform significantly better on writing, reading, attention, and working memory at ages 10-15. The results remain strong after controlling for differences across treatments and after controlling for baseline cognitive and noncognitive skills that could explain natural differences in ability across children. Columns (1) and (2) show that both math and reading scores are on average 0.4 and 0.3 standard deviations higher among children who select the later option in the Week 1 Immediate Decision

($p<0.01$). Columns (3)-(4) report results for our two primary measures of executive function, working memory and attention. Children who chose the later option in the Week 1 or Week 2 Immediate Decisions tend to perform 0.4-0.6 standard deviations better in working memory in future years ($p<0.01$). Children who chose the later option in the Week 2 Immediate Decision performed 0.4 standard deviations better on attention tasks ($p<0.10$). Lastly, column (5) shows a negative correlation between delaying in the Week 2 Immediate decision and discipline. However, this relationship is not systematically significant across all alternative explanations we consider in our robustness checks.

5.1 Robustness

Appendix Table A.11 reports robustness results for Section 5. In Table 7, derive our main outcome measures from exploratory factor analysis. To ensure that our results are not driven by estimated factor loadings and the way factor analysis imputes missing data, we replicate our results using outcomes where the measure is the average score across years (panel A). We also weight by the inverse probability of participating in the glow bracelet task conditional on CHECC group assignment (panel B). Lastly, to ensure our results are not driven by student compositional differences between days of the experiment, we restrict the sample to children who were present for all decision days (panel C).

Almost all results replicate for these alternative specifications. Across all panels, children who were able to delay in the Week 1 Immediate Decision at ages 5-10 performed significantly better in math, reading, and working memory assessments at ages 10-15. Children who delayed in the Week 2 Immediate Decision performed significantly better on working memory and attention measures. As mentioned above, the correlation between glow bracelet choices and discipline becomes insignificant when we use averages, rather than factors (panel A). We therefore cannot

rule out a null relationship between decisions in the glow bracelet task and discipline in the ensuing five years.

Table 7: Regression of Future Outcomes on Delay of Gratification Behavior

	Math	Reading	Working Memory	Attention	Discipline
	(1)	(2)	(3)	(4)	(5)
advance decision, week 1	0.193	0.214	-0.196	-0.074	-0.094
	[-0.030,0.409]	[-0.025,0.450]	[-0.464,0.057]	[-0.335,0.192]	[-0.228,0.049]
immediate decision, week 1	0.448	0.340	0.393	-0.049	-0.193
	[0.262,0.636]***	[0.146,0.539]***	[0.192,0.595]***	[-0.287,0.187]	[-0.424,0.029]
advance decision, week 2	0.022	0.156	-0.240	-0.319	0.054
	[-0.205,0.248]	[-0.063,0.374]	[-0.542,0.058]	[-0.607,-0.032]*	[-0.134,0.239]
immediate decision, week 2 (if no commit)	0.051	-0.136	0.643	0.448	-0.245
	[-0.265,0.365]	[-0.436,0.163]	[0.324,0.976]***	[0.060,0.838]*	[-0.478,-0.013]*
commit	0.120	0.014	0.350	0.227	-0.065
	[-0.175,0.418]	[-0.267,0.295]	[0.038,0.679]*	[-0.118,0.570]	[-0.267,0.127]
Observations	369	369	343	344	341
Child Gender & Race	Yes	Yes	Yes	Yes	Yes
Pre-assessment scores	Yes	Yes	Yes	Yes	Yes

Notes: Conventional standard errors in parentheses. Brackets show 90% CI after wild bootstrapping. Outcome variables are factors derived from exploratory factor analysis using measures taken after glow bracelet experiment. Outcomes in columns (1) and (2) are calculated from quarterly standardized math and reading assessment scores from Winter 2015 to Winter 2019, respectively. Outcomes in columns (3) and (4) are calculated from annual standardized tests of working memory and attention taken in 2016-2019, respectively. Outcome in column (5) derived from teacher disciplinary referrals. Controls: gender, race, baseline cognitive scores, and baseline noncognitive scores. Coefficients on preschool treatment, interaction terms between delay of gratification behavior and preschool treatment suppressed. Coefficients on dummies for missing data suppressed. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6. Conclusion

As Heckman (2008) argues, "skill begets skill". Human capital exhibits dynamic complementarities in that skills formed at early ages can improve the development of later skills. Our paper identifies the dynamic complementarity between time preferences and later reading, math, and executive function. To our knowledge, it is the first to use a natural field experiment to establish the importance of preschool in fostering these complementarities. Moreover, it provides evidence regarding the effectiveness of different curricula for developing patience. We find consistent and large impacts of the *Tools of the Mind* curriculum, which focuses on self-regulation, but little evidence of effectiveness for the *Literacy Express* curriculum, which focuses on traditional math and reading.

Children randomized to the *Tools* curriculum exhibit steadily high rates of patience throughout the experiment, in contrast to children in other groups, for whom patience declines from Week 1 to Week 2. We find little evidence that past behavior can explain this unanticipated trend, since patience, inconsistency, and impulsivity in Week 1 do not significantly differ across groups. Rather, social influences appear to play a role. In Week 2, children in the Control group tend to behave impatiently if more of their classmates behaved impatiently in the past, while children in the *Tools* group behave patiently independent of their classmates' past behavior. These results provide evidence for the claim in Falk et al. (2020) that interventions to help children delay gratification ought to consider the social environment in which preferences form, rather than teaching waiting strategies in a vacuum.

The *Tools* group is also the only group to use commitment strategically. We find that children from this curriculum are more likely to use commitment after behaving impulsively, whereas commitment use is uncorrelated with past behavior for all three other groups. For the *Tools* group, commitment is taken up by the very children who would benefit from it the most. This result aligns with the majority of prior work on commitment take-up among adults, which find higher commitment take-up for those who exhibited dynamic inconsistency in the past (Ariely and Wertenbroch, 2002; Ashraf et al., 2006; Kaur et al., 2010; Augenblick et al., 2015). Our findings suggest that the sophisticated use of external tools to meet long-term goals can be taught, and can originate as early as ages 5-10.

Finally, we find that children who delay even when the early option is paid immediately tend to perform better on academic and executive function measures in the ensuing five years. This result allows us to assess landmark studies in psychology which claim that delaying gratification at ages 3-5 somehow improves success in adulthood (Shoda et al., 1990; Mischel et al., 1972;

1989). We open up the black box behind this link by identifying the particular skills that are correlated with delaying gratification. Our results show that time preferences which help children delay gratification positively predict both cognitive skills like reading and math, as well as executive functions like attention and working memory. Our paper speaks to the importance of identifying how early interventions can equip children to further invest in their human capital and develop better skills at later ages.

References

Abadie, Alberto, Athey, Susan, Imbens, Guido, Wooldridge, Jeffrey M., 2022. When Should You Adjust Standard Errors for Clustering?

Alan, S. and Ertac, S., 2014. Good things come to those who (are taught how to) wait: Results from a randomized educational intervention on time preference.

Alan, S. and Ertac, S., 2015. Patience, self-control and the demand for commitment: Evidence from a large-scale field experiment. Journal of Economic Behavior & Organization, 115, pp.111-122.

Alan, S. and Ertac, S., 2018. Fostering patience in the classroom: Results from randomized educational intervention. Journal of Political Economy, 126(5), pp.1865-1911.

Alan, S., Ertac, S., and Gumus, I., 2021. Self-Control and the Demand for Commitment? Results from an Educational Intervention.

Andreoni, J., Kuhn, M.A., List, J.A., Samek, A., Sokal, K. and Sprenger, C., 2019. Toward an understanding of the development of time preferences: Evidence from field experiments. Journal of public economics, 177, p.104039.

Ariely, D., and Wertenbroch, K., 2002. Procrastination, Deadlines, and Performance: Self-Control by Precommitment. Psychological Science, 13, pp.219-224.

Ashraf, N., Karlan, D., Yin, W., 2006. Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines. The Quarterly Journal of Economics, 102, pp.1118-1139.

Augenblick, N., Niederle, M. and Sprenger, C., 2015. Working over time: Dynamic inconsistency in real effort tasks. The Quarterly Journal of Economics, 130(3), pp.1067-1115.

Augenblick, N. and Rabin, M., 2019. An experiment on time preference and misprediction in unpleasant tasks. Review of Economic Studies, 86(3), pp.941-975.

Belfield, C.R., Nores, M., Barnett, S. and Schweinhart, L., 2006. The high/scope perry preschool program cost–benefit analysis using data from the age-40 followup. Journal of Human resources, 41(1), pp.162-190.

Ben-Ner, A., List, J.A., Putterman, L. and Samek, A., 2017. Learned generosity? An artefactual field experiment with parents and their children. Journal of Economic Behavior & Organization, 143, pp.28-44.

Bodrova, E., & Leong, D. J. (2007). Tools of the Mind. (2nd ed.) Columbus, OH, Merrill/Prentice Hall

Blair, C., Raver, C.C., 2014. Closing the Achievement Gap through Modification of Neurocognitive and Neuroendocrine Function: Results from a Cluster Randomized Controlled Trial of an Innovative Approach to the Education of Children in Kindergarten. PLoS ONE 9(11): e112393. doi:10.1371/journal.pone.0112393

Cameron, A.C., Gelbach, J.B., Miller, D.L., 2008. Bootstrap-Based Improvements for Inference with Clustered Errors. Review of Economics and Statistics, 90(3), pp.414-427.

Campbell, F. A., Pungello, E. P., Miller-Johnson, S., Burchinal, M., & Ramey, C. T., 2001. The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. Developmental Psychology, 37(2), 231–242.

Campbell, F.A., Ramey, C.T., Pungello, E., Sparling, J. and Miller-Johnson, S., 2002. Early childhood education: Young adult outcomes from the Abecedarian Project. Applied developmental science, 6(1), pp.42-57.

Cappelen, A., List, J.A., Samek, A., Tungodden, B., 2020. The Effect of Early-Childhood Education on Social Preferences. Journal of Political Economy, 128(7), pp.2739-2758.

Carneiro, P., Heckman, J.J., 2003. Human Capital Policy (No. 9495). National Bureau of Economic Research.

Castillo, M., Jordan, J. L., and Petrie, R. (2018). Children's rationality, risk attitudes, and field behavior. European Economic Review, 102, pp.62–81.

Castillo, M., List, J.A., Petrie, R. and Samek, A., 2020. Detecting drivers of behavior at an early age: Evidence from a longitudinal field experiment (No. w28288). National Bureau of Economic Research.

Cowell, J.M., Samek, A., List, J. and Decety, J., 2015. The curious relation between theory of mind and sharing in preschool age children. PloS one, 10(2), p.e0117947.

Cox, J.C., List, J.A., Price, M., Sadiraj, V. and Samek, A., 2016. Moral costs and rational choice: Theory and experimental evidence (No. w22234). National Bureau of Economic Research.

Diamond, A., Barnett, W.S., Thomas, J., Munro, S., 2007. Preschool Program Improves Cognitive Control. Science, 318(5855), pp.1387-1388.

Diamond, A., Lee, C., Senften, P., Lam, A., Abbott, D., 2019. Randomized control trial of Tools of the Mind: Marked benefits to kindergarten children and their teachers. PLoS ONE 14(9): e0222447. https://doi.org/10.1371/journal. pone.0222447

Fairlie, R.W., 2005. An extension of the Blinder-Oaxaca decomposition technique to logit and probit models. Journal of Economic and Social Measurement 30: 305-316.

Falk, A., Kosse, F., Pinger, P., 2020. Re-Revisiting the Marshmallow Test: A Direct Comparison of Studies by Shoda, Mischel, and Peake (1990) and Watts, Duncan, and Quan (2018). Psychological Science, 31(1), pp100-104.

Fryer, R.G., Levitt, S.D. and List, J.A., 2015. Parental incentives and early childhood achievement: A field experiment in Chicago heights (No. w21477). National Bureau of Economic Research.

Hahn, R.A., Barnett, W.S., Knopf, J.A., et al. 2016. Early Childhood Education to Promote Health Equity: A Community Guide Systematic Review. Journal of public health management and practice, 22(5), pp.E1-E8.

Heckman, J.J. (2008). Schools, Skills, and Synapses. Economic Inquiry, 46(3), pp.289-324.

Heckman, J.J., Moon, S.H., Pinto, R., Savelyev, P.A. and Yavitz, A., 2010. The rate of return to the HighScope Perry Preschool Program. Journal of public Economics, 94(1-2), pp.114-128.

Heckman, J.J., Stixrud, J. and Urzua, S., 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. Journal of Labor economics, 24(3), pp.411-482.

Herrnstein, R.J., Loewenstein, G.F., Prelec, D. and Vaughan Jr, W., 1993. Utility maximization and melioration: Internalities in individual choice. Journal of behavioral decision making, 6(3), pp.149-185.

Houser, D., List, J.A., Piovesan, M., Samek, A.S. and Winter, J., 2015. On the origins of dishonesty: From parents to children (No. w20897). National Bureau of Economic Research.
Jann, B., 2008. The Blinder-Oaxaca decomposition for linear regression models. The Stata Journals 8(4):453-479.

Jones, S.M., Bouffard, S.M. Social and Emotional Learning in Schools: From Programs to Strategies. Social Policy Report, 26(4). Society for Research in Child Development.

Kaur, S., Kremer, M., Mullainathan, S., 2010. Self-Control and the Development of Work Arrangements. American Economic Review Papers and Proceedings, 92, pp.1606-1617.

Laibson, D., 1997. Golden eggs and hyperbolic discounting. The Quarterly Journal of Economics, 112(2), pp.443-478.

List, J.A., Petrie, R. and Samek, A., 2021. How experiments with children inform economics (No. w28825). National Bureau of Economic Research.

List, J.A. and Samak, A.C., 2013. Exploring the origins of charitable acts: Evidence from an artefactual field experiment with young children. Economics Letters, 118(3), pp.431-434.

List, J.A., Samek, A. and Suskind, D.L., 2018. Combining behavioral economics and field experiments to reimagine early childhood education. Behavioural Public Policy, 2(1), pp.1-21.

Lonigan, C.J., 2006. Impact of preschool literacy curricula: Results of a randomized evaluation in a public prekindergarten program.

Lonigan, C.J., Faver, J.M., Clancy-Menchetti, J., Phillips, B.M., 2005. Promoting the development of preschool children's emergent literacy skills: A randomized evaluation of a literacy-focused curriculum and two professional development models.

Luhrmann, M., Serra-Garcia, M. and Winter, J., 2018. The impact of financial education on adolescents' intertemporal choices. American Economic Journal: Economic Policy, 10(3), pp.309-32.

Masse, L.N. and Barnett, W.S., 2002. A benefit-cost analysis of the Abecedarian Early Childhood Intervention. Cost-effectiveness and educational policy, 2002, p.157.

Mischel, W., Ebbesen, E.B. and Raskoff Zeiss, A., 1972. Cognitive and attentional mechanisms in delay of gratification. Journal of personality and social psychology, 21(2), p.204-218.

Mischel, W., Shoda, Y. and Rodriguez, M.L., 1989. Delay of gratification in children. Science, 244(4907), pp.933-938.

Nores, M., Belfield, C.R., Barnett, W.S. and Schweinhart, L., 2005. Updating the economic impacts of the High/Scope Perry Preschool program. Educational Evaluation and Policy Analysis, 27(3), pp.245-261.

O'Donoghue, T. and Rabin, M., 1999. Doing it now or later. American economic review, 89(1), pp.103-124.

O'Donoghue, T. and Rabin, M., 2006. Optimal sin taxes. Journal of Public Economics, 90(10-11), pp.1825-1849.

Ramon, I., Chattopadhyay, S.K., Barnett, W.S., Hahn, R.A., 2018. Early Childhood Education to Promote Health Equity: A Community Guide Economic Review. J Public Health Manag Pract., 24(1), pp.e8-e15.

Rao, G., 2019. Familiarity does not breed contempt. The American Economic Review, 109(3), pp.774-809.

Roodman, D., MacKinnon, J.G., Nielsen, M.O., Webb, M.D., 2019. Fast and wild: Bootstrap inference in Stata using boottest. The Stata Journal 19(1), 4-60.

Sadoff, S. and Samek, A., 2019. Can interventions affect commitment demand? A field experiment on food choice. Journal of Economic Behavior & Organization, 158, pp.90-109

Sadoff, S., Samek, A. and Sprenger, C., 2020. Dynamic inconsistency in food choice: Experimental evidence from two food deserts. The Review of Economic Studies, 87(4), pp.1954-1988.

Samek, A., 2013. Is there a gender gap in preschoolers' competitiveness? an experiment in the US. Journal of Economic Behavior and Organization, 92.

Samek, A. and Sheremeta, R.M., 2017. Selective recognition: How to recognize donors to increase charitable giving. Economic Inquiry, 55(3), pp.1489-1496.

Shoda, Y., Mischel, W. and Peake, P.K., 1990. Predicting adolescent cognitive and self-regulatory competencies from preschool delay of gratification: Identifying diagnostic conditions. Developmental psychology, 26(6), p.978-986.

Strotz, R.H., 1955. Myopia and inconsistency in dynamic utility maximization. The review of economic studies, 23(3), pp.165-180.

Thaler, R.H. and Shefrin, H.M., 1981. An economic theory of self-control. Journal of political Economy, 89(2), pp.392-406.

Watts, T. W., Duncan, G. J., & Quan, H. (2018). Revisiting the marshmallow test: A conceptual replication investigating links between early delay of gratification and later outcomes. Psychological Science, 29, pp.1159–1177.

Wu, C.F.G., 1986. Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis. Annals of Statistics 14, 1261-129.

Appendix A: Additional Tables and Figures

Table A.1: Income bracket and elementary school, by CHECC Treatment

		Years 1-2			Years 3-4	
	All	Control	Tools	Lit	Control	CogX
			A: Income Bracket			
income: 0-25K	0.322	0.403	0.333	0.385	0.179	0.356
	(0.024)	(0.042)	(0.076)	(0.079)	(0.036)	(0.063)
income: 26-60K	0.126	0.137	0.256	0.154	0.089	0.068
	(0.017)	(0.029)	(0.071)	(0.059)	(0.027)	(0.033)
income: 61K or more	0.010	0.014	0.026	0.000	0.000	0.017
	(0.005)	(0.010)	(0.026)	(0.000)	(0.000)	(0.017)
			B: Elementary School			
School A	0.206	0.194	0.128	0.282	0.205	0.237
	(0.021)	(0.034)	(0.054)	(0.073)	(0.038)	(0.056)
School B	0.121	0.151	0.205	0.103	0.098	0.051
	(0.017)	(0.030)	(0.066)	(0.049)	(0.028)	(0.029)
School C	0.085	0.108	0.077	0.051	0.062	0.102
	(0.014)	(0.026)	(0.043)	(0.036)	(0.023)	(0.040)
School D	0.064	0.072	0.026	0.051	0.071	0.068
	(0.012)	(0.022)	(0.026)	(0.036)	(0.024)	(0.033)
School E	0.041	0.029	0.077	0.000	0.071	0.017
	(0.010)	(0.014)	(0.043)	(0.000)	(0.024)	(0.017)
School F	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
School G	0.067	0.058	0.026	0.103	0.089	0.051
	(0.013)	(0.020)	(0.026)	(0.049)	(0.027)	(0.029)
Observations	388	139	39	39	112	59

Notes: Standard errors in parentheses. Stars denote significance relative to control after Sidak correction for multiple hypothesis testing. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2: Proportion Inconsistent & Impulsive

	Week 1			Week 2			
	Patient	Inconsistent	Impulsive	Patient	Inconsistent	Impulsive	Commit
All	(80.8%)	(21.4%)	(15.8%)	(75.7%)	(18.7%)	(10.3%)	(51.2%)
	298/369	77/360	57/360	256/338	29/155	16/155	173/338
Control, Years 1-2	(87.5%)	(15.2%)	(12.1%)	(73.8%)	(23.5%)	(9.8%)	(55.7%)
	119/136	20/132	16/132	90/122	12/51	5/51	68/122
Tools	(88.2%)	(11.8%)	(8.8%)	(97.1%)	(14.3%)	(7.1%)	(55.9%)
	30/34	4/34	3/34	33/34	2/14	1/14	19/34
Literacy	(86.1%)	(14.3%)	(14.3%)	(78.8%)	(0.0%)	(0.0%)	(57.6%)
	31/36	5/35	5/35	26/33	0/13	0/13	19/33
Control, Years 3-4	(73.1%)	(32.1%)	(21.7%)	(69.4%)	(22.4%)	(14.3%)	(49.0%)
	79/108	34/106	23/106	68/98	11/49	7/49	48/98
CogX	(70.9%)	(26.4%)	(18.9%)	(76.5%)	(14.3%)	(10.7%)	(37.3%)
	39/55	14/53	10/53	39/51	4/28	3/28	19/51

Notes: Proportion patient, inconsistent, and impulsive by treatment and week.

Table A.3: Regression of Immediate Decision (exc. Committers) on CHECC Treatment

	Chose Later Option			
	(1)	(2)	(3)	(4)
Week 2	-0.039	-0.048	-0.043	-0.031
	[-0.130,0.050]	[-0.139,0.042]	[-0.135,0.048]	[-0.121,0.055]
Tools	-0.076	-0.081	-0.094	-0.046
	[-0.206,0.059]	[-0.210,0.056]	[-0.229,0.047]	[-0.168,0.077]
Tools × Week 2	0.168	0.172	0.169	0.118
	[0.022,0.310]*	[0.025,0.314]*	[0.016,0.312]*	[-0.016,0.244]
Literacy	-0.137	-0.147	-0.175	-0.097
	[-0.272,0.001]	[-0.286,-0.005]*	[-0.313,-0.033]*	[-0.222,0.028]
Literacy × Week 2	0.146	0.143	0.138	0.132
	[-0.073,0.372]	[-0.076,0.369]	[-0.074,0.363]	[-0.092,0.369]
CogX	0.018	0.027	0.045	0.053
	[-0.117,0.152]	[-0.112,0.166]	[-0.098,0.188]	[-0.072,0.178]
CogX × Week 2	0.189	0.181	0.190	0.187
	[0.003,0.381]*	[-0.007,0.373]	[0.004,0.378]*	[-0.010,0.384]
Observations	516	516	516	516
Child Gender & Race		Yes	Yes	Yes
Pre-assessment scores			Yes	Yes
Parental Controls				Yes

Notes: All regressions control for year of CHECC participation and child age. Parental controls include age, HS graduate status, and full-time employment of mother and father. Dummies for missing variables suppressed. Standard errors clustered by child. 90% confidence intervals use wild bootstrap-t method from Cameron, Gelbach, and Miller (2008). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Regression of Patience in Week 2 on Impulsive Behavior in Week 1

	Patience in Week 2			
	(1)	(2)	(3)	(4)
Inconsistent in week 1	-0.073	-0.063	-0.054	-0.024
	[-0.210,0.058]	[-0.195,0.068]	[-0.187,0.079]	[-0.159,0.108]
PK Tools	0.198	0.181	0.202	0.205
	[0.105,0.288]***	[0.086,0.272]***	[0.102,0.298]***	[0.093,0.306]***
Inconsistent in week 1 × PK Tools	0.110	0.081	0.085	0.164
	[-0.079,0.301]	[-0.141,0.295]	[-0.141,0.306]	[-0.064,0.403]
PK Literacy	0.049	0.027	0.044	0.061
	[-0.101,0.198]	[-0.124,0.176]	[-0.118,0.207]	[-0.099,0.216]
Inconsistent in week 1 × PK Lit	-0.113	-0.109	-0.145	-0.079
	[-0.723,0.544]	[-0.756,0.595]	[-0.789,0.559]	[-0.777,0.640]
PK CogX	0.073	0.075	0.089	0.088
	[-0.083,0.225]	[-0.078,0.227]	[-0.076,0.251]	[-0.074,0.251]
Inconsistent in week 1 × PK CogX	-0.007	-0.005	-0.011	0.010
	[-0.297,0.283]	[-0.295,0.295]	[-0.303,0.279]	[-0.294,0.312]
Observations	335	335	335	335
Child Gender & Race		Yes	Yes	Yes
Pre-assessment scores			Yes	Yes
Parental Controls				Yes

Notes: All regressions control for year of CHECC participation and child age. Parental controls include age, high school graduate status, and full-time employment of mother. Dummies for missing variables suppressed. Standard errors clustered by child. 90% confidence intervals use wild bootstrap-t method from Cameron, Gelbach, and Miller (2008). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Oaxaca decomposition on % classmates with 2 bracelets

	Choose Later Option, Week 2 Advance Choice *Literacy* vs. Control			
	Overall			
	(1)	(2)	(3)	(4)
Control	0.757***			
	(0.0563)			
Literacy	0.828***			
	(0.0768)			
Control -*Literacy*	-0.0708			
	(0.0880)			
Endowments	-0.0996			
	(0.119)			
Coefficients	-0.172			
	(0.115)			
Interaction	0.200			
	(0.137)			
Decomposition:		Endowments	Coefficient	Interaction
% classmates with 2 bracelets		0.00311	0.494	0.0350
		(0.0129)	(0.255)	(0.0522)
child age		-0.0491	-1.882	0.0331
		(0.0765)	(3.138)	(0.0668)
female		-0.0284	-0.102	0.0203
		(0.0349)	(0.104)	(0.0289)
black		0.0477	-0.0868	-0.0436
		(0.0575)	(0.0741)	(0.0570)
hispanic		-0.0429	-0.735**	0.0421
		(0.0953)	(0.278)	(0.0941)
baseline cog (std)		-0.00162	0.00534	0.00446
		(0.0126)	(0.0175)	(0.0148)
baseline noncog (std)		0.00604	0.00324	-0.00794
		(0.0122)	(0.0102)	(0.0149)
mother's age		-0.121	-0.706*	0.125
		(0.117)	(0.343)	(0.127)
mother is HS grad		-0.00360	-0.0987	0.00630
		(0.0141)	(0.0860)	(0.0242)
mother works full-time		-0.0195	0.0222	0.0445
		(0.0393)	(0.0260)	(0.0467)
N		140		

Standard errors clustered by classroom in parentheses. All regressions additionally control for years of CHECC participation and dummies for missing variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.6: Oaxaca decomposition on % classmates with 2 bracelets

	Choose Later Option, Week 2 Advance Choice *CogX* vs. Control			
	Overall			
	(1)	(2)	(3)	(4)
Control	0.674***			
	(0.0428)			
CogX	0.756***			
	(0.0567)			
Control - *CogX*	-0.0814			
	(0.0765)			
Endowments	0.208			
	(0.197)			
Coefficients	-0.0945			
	(0.107)			
Interaction	-0.195			
	(0.213)			
Decomposition:		Endowments	Coefficient	Interaction
% classmates with 2 bracelets		-0.00297	0.0246	0.00135
		(0.00842)	(0.151)	(0.00857)
child age		-0.0269	1.776	0.0284
		(0.0241)	(1.333)	(0.0287)
female		0.00916	-0.167*	-0.00919
		(0.0297)	(0.0662)	(0.0299)
black		-0.0300	0.0665	0.0232
		(0.0353)	(0.0927)	(0.0398)
hispanic		0.0539	0.172	-0.0329
		(0.0469)	(0.386)	(0.0749)
baseline cog (std)		-0.0212	-0.0804*	0.0454
		(0.0165)	(0.0384)	(0.0300)
baseline noncog (std)		-0.0147	0.0288	0.0143
		(0.0299)	(0.0268)	(0.0298)
mother's age		0.00238	-0.0137	-0.000979
		(0.0167)	(0.187)	(0.0145)
mother is HS grad		-0.00101	0.0580	0.000652
		0.0377)	(0.0648)	(0.0243)
mother works full-time		-0.000644	0.0612	0.000687
		(0.0415)	(0.0453)	(0.0442)
N		123		

Standard errors clustered by classroom in parentheses. All regressions additionally control for years of CHECC participation and dummies for missing variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.7: Regression of commitment on impulsive behavior in Week 1

	Commitment in Week 2			
	(1)	(2)	(3)	(4)
Advance decision, week 1	0.011	0.004	-0.004	-0.006
	[-0.151,0.173]	[-0.152,0.163]	[-0.167,0.158]	[-0.171,0.153]
Immediate decision, week 1	-0.066	-0.021	-0.031	-0.043
	[-0.216,0.085]	[-0.169,0.130]	[-0.181,0.119]	[-0.207,0.124]
Advance decision, week 2	0.088	0.077	0.084	0.078
	[-0.046,0.220]	[-0.059,0.208]	[-0.053,0.218]	[-0.058,0.212]
PK Tools	0.088	0.077	0.084	0.078
	[.,.]	[.,.]	[.,.]	[.,.]
PK Literacy	0.088	0.077	0.084	0.078
	[-0.796,0.641]	[-0.713,0.439]	[-0.705,0.412]	[-0.681,0.465]
PK CogX	0.088	0.077	0.084	0.078
	[-0.360,0.394]	[-0.386,0.383]	[-0.351,0.356]	[-0.336,0.379]
Advance decision, week 1 × PK Tools	0.527	0.587	0.602	0.570
	[-0.732,0.920]	[-0.288,0.972]	[-0.198,0.984]	[-0.182,1.017]
Advance decision, week 1 × PK Lit	0.149	0.172	0.172	0.165
	[-0.498,0.710]	[-0.403,0.684]	[-0.419,0.671]	[-0.426,0.682]
Advance decision, week 1 × PK CogX	-0.094	-0.050	-0.047	-0.022
	[-0.462,0.239]	[-0.420,0.285]	[-0.415,0.284]	[-0.404,0.317]
Immediate decision, week 1 × PK Tools	-0.354	-0.412	-0.400	-0.444
	[-0.572,-0.156]**	[-0.617,-0.191]**	[-0.608,-0.176]**	[-0.690,-0.184]**
Immediate decision, week 1 × PK Lit	0.283	0.227	0.217	0.188
	[-0.351,0.745]	[-0.354,0.709]	[-0.352,0.711]	[-0.413,0.731]
Immediate decision, week 1 × PK CogX	0.167	0.076	0.080	0.052
	[-0.143,0.472]	[-0.239,0.399]	[-0.230,0.406]	[-0.257,0.382]
Advance decision, week 2 × PK Tools	0.465	0.384	0.403	0.446
	[-4.998,5.317]	[-6.758,6.139]	[-7.466,8.244]	[-5.639,5.842]
Advance decision, week 2 × PK Lit	-0.260	-0.238	-0.252	-0.238
	[-0.677,0.170]	[-0.704,0.226]	[-0.728,0.219]	[-0.735,0.241]
Advance decision, week 2 × PK CogX	-0.229	-0.163	-0.171	-0.187
	[-0.560,0.119]	[-0.513,0.190]	[-0.523,0.183]	[-0.535,0.169]
Observations	335	335	335	335
Child Gender & Race		Yes	Yes	Yes
Pre-assessment scores			Yes	Yes
Parental Controls				Yes

Notes: All regressions control for year of CHECC participation and child age. Parental controls include age, high school graduate status, and full-time employment of mother and father. Dummies for missing variables suppressed. Standard errors clustered by child. 90% confidence intervals use wild bootstrap-t method from Cameron, Gelbach, and Miller (2008). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8: Regression of Advance Decision on CHECC Treatment, Robustness Checks

	Chose Later Option, Advance Decision			
	(1)	(2)	(3)	(4)
Week 2	-0.095	-0.185	-0.085	-0.184
	(0.036)**	(0.079)**	(0.038)**	(0.083)**
	[-0.151,-0.039]***	[-0.322,-0.047]**	[-0.147,-0.027]**	[-0.341,-0.033]**
Tools	-0.020	-0.049	0.020	-0.022
	(0.066)	(0.076)	(0.066)	(0.080)
	[-0.137,0.098]	[-0.179,0.080]	[-0.090,0.132]	[-0.164,0.120]
Tools × Week 2	0.183	0.279	0.148	0.259
	(0.074)**	(0.099)**	(0.073)*	(0.102)**
	[0.059,0.309]**	[0.103,0.451]***	[0.034,0.265]**	[0.083,0.432]**
Literacy	-0.032	-0.045	-0.030	-0.047
	(0.067)	(0.070)	(0.077)	(0.081)
	[-0.148,0.083]	[-0.163,0.079]	[-0.159,0.108]	[-0.183,0.092]
Literacy × Week 2	0.013	0.040	0.022	0.090
	(0.088)	(0.120)	(0.094)	(0.117)
	[-0.136,0.165]	[-0.178,0.243]	[-0.138,0.177]	[-0.121,0.297]
CogX	-0.037	-0.032	0.009	-0.011
	(0.074)	(0.074)	(0.080)	(0.082)
	[-0.160,0.085]	[-0.152,0.085]	[-0.126,0.140]	[-0.153,0.127]
CogX × Week 2	0.153	0.202	0.130	0.203
	(0.087)*	(0.112)*	(0.092)	(0.120)
	[0.010,0.296]*	[0.015,0.388]*	[-0.020,0.283]	[0.003,0.410]*
Observations	707	707	633	633
Inverse probability weights		Yes		Yes
Present all days			Yes	Yes

Notes: All regressions control for age, race, gender, baseline cognitive scores, baseline noncognitive scores, and year of CHECC participation Columns (2) and (4) use inverse weights for the probability of completing the glow bracelet task conditional on group. Columns (3) and (4) restrict the sample to only children who were present for all decisions in the glow bracelet task. Dummies for missing variables suppressed. Conventional standard errors, clustered by child, in parentheses. 90% confidence intervals use wild bootstrap-t method from Cameron, Gelbach, and Miller (2008). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: Regression of Immediate Decision (inc. Committers) on CHECC Treatment, Robustness Checks

	Chose Later Option, Immediate Decision (inc. Committers)			
	(1)	(2)	(3)	(4)
Week 2	0.002	-0.107	-0.000	-0.116
	(0.039)	(0.089)	(0.040)	(0.091)
	[-0.061,0.067]	[-0.271,0.058]	[-0.064,0.064]	[-0.289,0.054]
Tools	-0.044	-0.059	-0.023	-0.048
	(0.078)	(0.084)	(0.080)	(0.088)
	[-0.176,0.088]	[-0.204,0.084]	[-0.154,0.113]	[-0.190,0.102]
Tools × Week 2	0.146	0.220	0.125	0.211
	(0.073)*	(0.102)**	(0.072)*	(0.104)*
	[0.021,0.268]*	[0.041,0.406]**	[0.008,0.241]*	[0.028,0.389]*
Literacy	-0.130	-0.137	-0.105	-0.141
	(0.082)	(0.091)	(0.089)	(0.099)
	[-0.270,0.014]	[-0.291,0.008]	[-0.254,0.050]	[-0.315,0.031]
Literacy × Week 2	0.045	0.076	0.034	0.124
	(0.093)	(0.126)	(0.088)	(0.112)
	[-0.106,0.195]	[-0.126,0.282]	[-0.120,0.182]	[-0.077,0.324]
CogX	0.021	0.069	0.015	0.046
	(0.081)	(0.090)	(0.089)	(0.100)
	[-0.116,0.154]	[-0.085,0.221]	[-0.142,0.173]	[-0.128,0.218]
CogX × Week 2	0.118	0.225	0.133	0.260
	(0.095)	(0.118)*	(0.102)	(0.125)*
	[-0.050,0.287]	[0.020,0.431]*	[-0.029,0.300]	[0.049,0.472]**
Observations	689	689	634	634
Inverse probability weights		Yes		Yes
Present all days			Yes	Yes

Notes: All regressions control for age, race, gender, baseline cognitive scores, baseline noncognitive scores, and year of CHECC participation Columns (2) and (4) use inverse weights for the probability of completing the glow bracelet task conditional on group. Columns (3) and (4) restrict the sample to only children who were present for all decisions in the glow bracelet task. Dummies for missing variables suppressed. Conventional standard errors, clustered by child, in parentheses. 90% confidence intervals use wild bootstrap-t method from Cameron, Gelbach, and Miller (2008). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.10: Regression of commit on impulsive behavior in Week 1, robustness checks

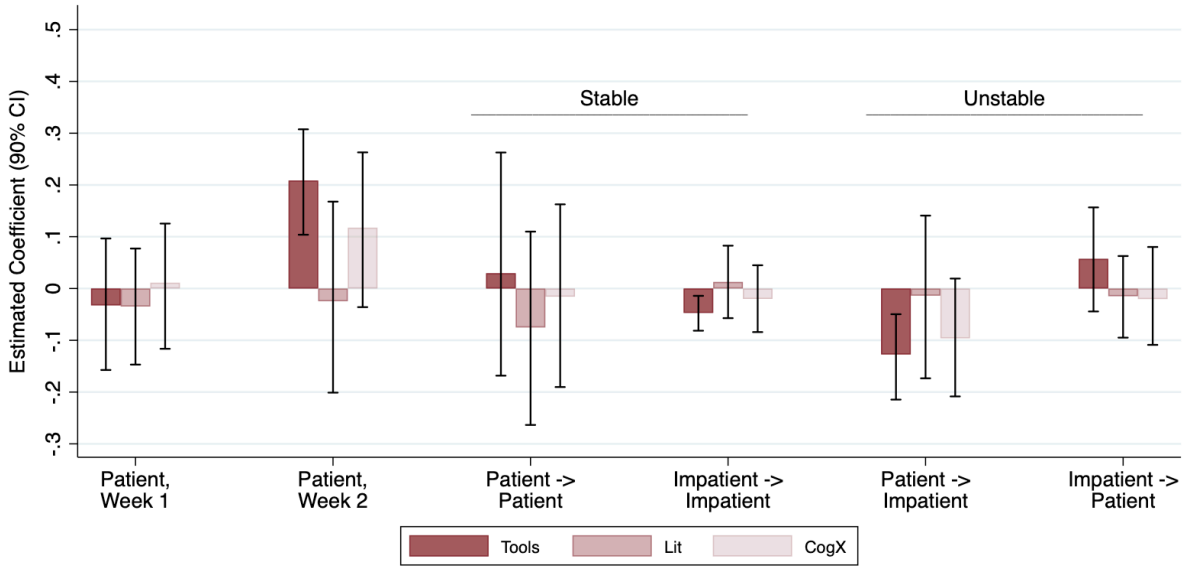
	Commitment in Week 2			
	(1)	(2)	(3)	(4)
Impulsivity in week 1	0.022	-0.036	0.002	-0.048
	(0.102)	(0.113)	(0.104)	(0.116)
	[-0.154,0.197]	[-0.242,0.176]	[-0.183,0.181]	[-0.263,0.166]
PK Tools	-0.152	-0.175	-0.108	-0.137
	(0.108)	(0.120)	(0.112)	(0.125)
	[-0.334,0.026]	[-0.376,0.029]	[-0.292,0.081]	[-0.347,0.074]
PK Literacy	-0.035	0.034	-0.036	0.016
	(0.102)	(0.103)	(0.107)	(0.109)
	[-0.203,0.154]	[-0.144,0.205]	[-0.220,0.159]	[-0.172,0.203]
PK CogX	-0.036	-0.066	0.004	-0.030
	(0.103)	(0.112)	(0.110)	(0.118)
	[-0.201,0.134]	[-0.251,0.134]	[-0.188,0.190]	[-0.243,0.189]
Impulsive in week 1 × PK Tools	0.492	0.603	0.512	0.607
	(0.149)***	(0.169)***	(0.157)***	(0.178)***
	[0.153,0.683]*	[0.116,0.907]*	[0.275,0.731]*	[0.209,0.859]*
Impulsive in week 1 × PK Lit	-0.109	-0.192	-0.139	-0.193
	(0.226)	(0.249)	(0.301)	(0.322)
	[-0.626,0.472]	[-0.700,0.756]	[-0.892,1.820]	[-0.950,3.778]
Impulsive in week 1 × PK CogX	-0.221	-0.006	-0.260	-0.043
	(0.179)	(0.201)	(0.183)	(0.202)
	[-0.543,0.104]	[-0.472,0.369]	[-0.599,0.072]	[-0.514,0.337]
Observations	335	335	316	316
Inverse probability weights		Yes		Yes
Present all days			Yes	Yes

Notes: All regressions control for age, race, gender, baseline cognitive scores, baseline noncognitive scores, and year of CHECC participation Columns (2) and (4) use inverse weights for the probability of completing the glow bracelet task conditional on group. Columns (3) and (4) restrict the sample to only children who were present for all decisions in the glow bracelet task. Dummies for missing variables suppressed. Conventional standard errors, clustered by child, in parentheses. 90% confidence intervals use wild bootstrap-t method from Cameron, Gelbach, and Miller (2008). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

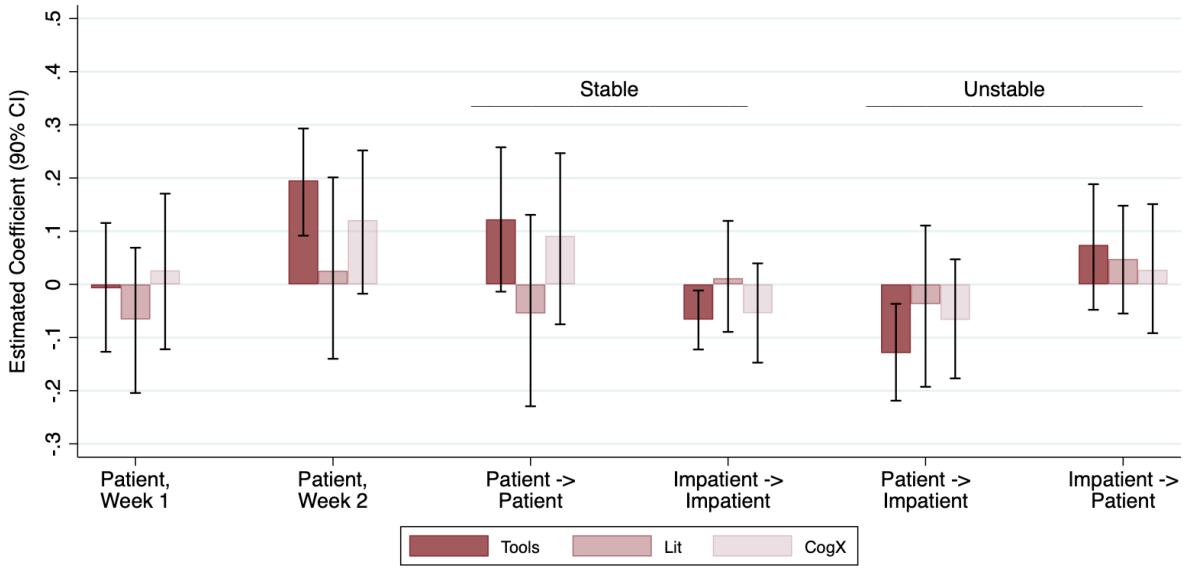
Table A.11: Regression of Future Outcomes on Delay of Gratification Behavior, Robustness Checks

	Math	Reading	Working Memory	Attention	Discipline
	(1)	(2)	(3)	(4)	(5)
	A. Outcomes: average score across years				
advance decision, week 1	0.636	2.089	-3.195	0.432	0.147
	[-2.956,4.094]	[-1.585,5.668]	[-6.343,-0.222]*	[-2.348,3.196]	[-0.134,0.407]
immediate decision, week 1	6.655	4.619	4.519	1.132	0.446
	[3.754,9.614]***	[1.554,7.749]**	[1.902,7.134]***	[-1.491,3.718]	[-0.013,0.928]
advance decision, week 2	0.425	2.989	-4.059	-3.627	-0.150
	[-3.085,3.957]	[-0.530,6.502]	[-7.656,-0.512]*	[-6.954,-0.263]*	[-0.522,0.231]
immediate decision, week 2 (if no commit)	1.314	-2.447	8.721	4.946	0.366
	[-4.519,6.976]	[-8.184,3.182]	[4.453,13.026]***	[0.485,9.391]*	[-0.128,0.865]
commit	3.680	1.523	4.276	2.925	0.096
	[-1.589,8.846]	[-3.676,6.639]	[-0.014,8.709]	[-0.955,6.794]	[-0.317,0.535]
Observations	369	369	343	344	341
	B. Inverse probability weights				
advance decision, week 1	0.194	0.208	-0.191	-0.085	-0.099
	[-0.029,0.408]	[-0.032,0.444]	[-0.459,0.062]	[-0.344,0.177]	[-0.222,0.031]
immediate decision, week 1	0.452	0.339	0.393	-0.025	-0.179
	[0.266,0.639]***	[0.145,0.537]***	[0.194,0.593]***	[-0.261,0.209]	[-0.381,0.016]
advance decision, week 2	0.008	0.143	-0.246	-0.303	0.048
	[-0.219,0.234]	[-0.076,0.360]	[-0.550,0.055]	[-0.590,-0.016]*	[-0.140,0.232]
immediate decision, week 2 (if no commit)	0.066	-0.116	0.630	0.499	-0.231
	[-0.249,0.379]	[-0.412,0.181]	[0.310,0.964]***	[0.125,0.879]**	[-0.460,-0.003]*
commit	0.135	0.034	0.337	0.269	-0.047
	[-0.160,0.432]	[-0.244,0.312]	[0.022,0.668]*	[-0.059,0.600]	[-0.245,0.142]
Observations	369	369	343	344	341
	C. Children present for all decisions				
advance decision, week 1	0.145	0.160	-0.178	-0.094	-0.117
	[-0.092,0.379]	[-0.101,0.417]	[-0.462,0.092]	[-0.393,0.202]	[-0.285,0.056]
immediate decision, week 1	0.467	0.335	0.385	0.097	-0.062
	[0.275,0.661]***	[0.135,0.538]***	[0.173,0.598]***	[-0.159,0.354]	[-0.296,0.173]
advance decision, week 2	-0.014	0.171	-0.249	-0.326	0.015
	[-0.249,0.219]	[-0.056,0.393]	[-0.573,0.070]	[-0.622,-0.031]*	[-0.330,0.359]
immediate decision, week 2 (if no commit)	0.150	-0.045	0.599	0.581	-0.419
	[-0.182,0.486]	[-0.359,0.274]	[0.254,0.947]***	[0.203,0.975]**	[-0.802,-0.048]*
commit	0.202	0.106	0.305	0.406	-0.099
	[-0.113,0.522]	[-0.193,0.410]	[-0.039,0.653]	[0.074,0.757]**	[-0.383,0.165]
Observations	307	307	285	284	282
Child Gender & Race	Yes	Yes	Yes	Yes	Yes
Pre-assessment scores	Yes	Yes	Yes	Yes	Yes

Notes: In panel A, the outcome is the average measure taken across years, rather than the factor derived from exploratory factor analysis. In panels B and C, the outcome is the factor derived from exploratory factor analysis. Panel B incorporates inverse probability weights to account for potential differences in the likelihood of participating in the glow bracelet task conditional on CHECC group assignment. Panel C restricts the sample to only children who were present for all decisions.
Conventional standard errors in parentheses. Brackets show 90% CI after wild bootstrapping. Outcome variables are factors derived from exploratory factor analysis using measures taken after glow bracelet experiment. Outcomes in columns (1) and (2) are calculated from quarterly standardized math and reading assessment scores from Winter 2015 to Winter 2019, respectively. Outcomes in columns (3) and (4) are calculated from annual standardized tests of working memory and attention taken in 2016-2019, respectively. Outcome in column (5) derived from teacher disciplinary referrals. Controls: gender, race, baseline cognitive scores, and baseline noncognitive scores. Coefficients on preschool treatment, interaction terms between delay of gratification behavior and preschool treatment suppressed. Coefficients on dummies for missing data suppressed. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



a.



b.

Appendix Figure A.1: Difference in Advance Choices relative to Control group.

Notes: Panel a incorporates inverse probability weights to adjust for possibility that CHECC preschool assignment affected likelihood of enrolling in Chicago Heights school district. Panel b restricts the sample to only children present for all decisions in the glow bracelet task. Controls: age, race, gender, year of CHECC participation, baseline cognitive and noncognitive scores.

Appendix B

Experiment Scripts

B.0.1 Assent (Day 1 only)

Today, I am asking you all to do an activity with me where I will ask you some questions about choices you want to make. You will get some glow bracelets for doing the activity. We will keep your answers to look at later, to help us learn more about how kids decide. You do not have to do this activity. It is up to you if you want to do the activity or not. You can say no now or you can even change your mind later. No one will be upset with you if you don't want to do this activity.

Does anyone have any questions for me? You can ask me questions at any time. Will you do the activity with me?

B.0.2 Week 1 Day 1

Thanks for doing the activity. First, we are going to sit quietly. It is very important that we do not talk during the activity. If anyone talks, they will have to go back to class and will not get any prizes. Can we practice being quiet? [wait for kids to settle down]. Here are the glow bracelets that we will play with. (Walk around with tray of glow bracelets.)

In the activity, you are going to decide WHEN you want to get glow bracelets. When you get the glow bracelets, they will glow for 1 day. But after 1 day they will stop glowing. Today you will decide whether you want 1 glow bracelet TOMORROW - that's sooner, OR 2 glow bracelets the DAY AFTER TOMORROW - that's later. You will get your glow bracelet in an envelope at the end of the day that you decide to get it, and get a blank envelope on the other day. The choice you make is up to you. You will only get glow bracelets if you follow the rules for the activity sheet. Here is the activity sheet you will have (Show poster 1). Here are the rules:

- If you want 1 glow bracelet TOMORROW - that's SOONER, you should circle this choice. See this shows ONE glow bracelet. It also shows that you will have to sleep for ONE night before tomorrow comes. Tomorrow is TUESDAY/THURSDAY. (point to the 1 glow bracelet sheet as well as each component).

- If you want 2 glow bracelets the DAY AFTER TOMORROW - that's LATER, circle this choice. See this shows TWO glow bracelets. It also shows that you will have to sleep for TWO nights before the day after tomorrow comes. The day after tomorrow IS WEDNESDAY/FRIDAY. (Point to the 2 glow bracelets sheet and each component).



Figure B1: Poster 1 and response sheets: Monday cohort received top panel and Wednesday cohort received bottom panel

Doing the Practice Book

Before we start the activity, you are each going to fill out a practice book to make sure you understand the rules. Please remember to try your best! Do not look at other people's work and no talking! Only turn the pages when I tell you to. I will read out loud and you should follow along. You should circle your answers. (Pass out the practice book and pencils and read out loud.) Okay let's start. Okay please open the practice book. Remember its okay if you don't know the answer. Just try your best.

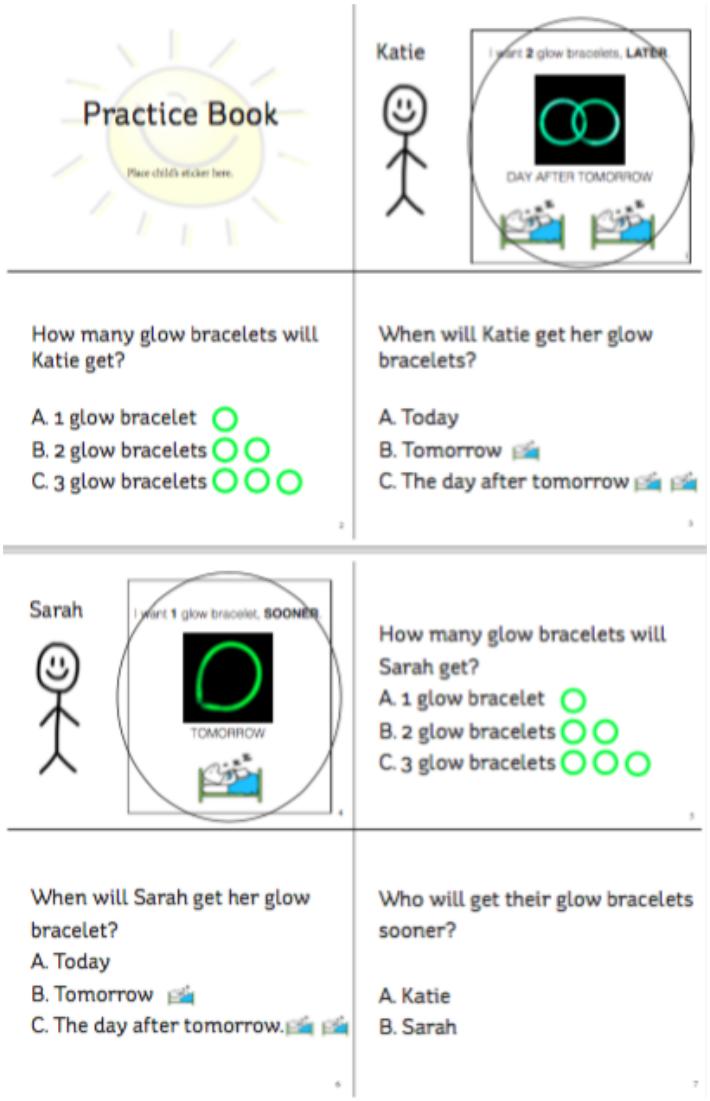


Figure B2: Practice Book Screenshots

- This is Katie. Katie played the same game that you are going to play. She decided to circle THIS choice, 2 glow bracelets, LATER. LATER is the DAY AFTER TOMORROW, WEDNESDAY/ FRIDAY. (Say 'Ok, turn the page' after each page you read.)

  - The practice book asks, How many bracelets will Katie get? 1, 2 or 3? (Point to each answer.) Circle your answer now. Ok, turn the page.
  - The practice book asks, When will Katie get her bracelets? Today, tomorrow, or the day after tomorrow? (Point to each answer) Circle your answer now. Ok, turn the page.

- Now let's do another one. This is Sarah. Sarah played the same game that you are going to play. She decided to circle THIS choice, 1 glow bracelet, SOONER. SOONER is tomorrow. Ok, turn the page.
  - The practice book asks, How many bracelets will Sarah get? 1, 2, or 3? (Point to each answer.) Circle your answer now. Ok, turn the page.
  - The practice book asks, When will Sarah get her bracelets? Today, tomorrow, or the day after tomorrow? (Point to each answer.) Circle your answer now. Ok, turn the page.
  - The practice book asks, Who will get their bracelet sooner? Remember, this was Katie's choice (Point back to Katie picture). This was Sarah's choice (point back to Sarah Picture). So who will get their bracelet sooner, Katie, or Sarah? (Point to each answer). Circle your answer now.

Great job! Now I will collect your practice books and we can start the activity.

Doing the Activity

(go back to poster 1) Now remember,

- If you want 1 glow bracelet TOMORROW - that's SOONER, you should circle this choice (Point to the 1 glow bracelet sheet).
- If you want 2 glow bracelets the DAY AFTER TOMORROW - that's LATER, circle this choice (Point to the 2 glow bracelets sheet)

Here are your activity sheets. Remember to make your decision quietly, keep your eyes on your paper and do not talk to anyone. If you have a question, raise your hand and I will come to you. Don't start until I say GO. [pass out decision sheets]. Whatever you choose, we will give you the glow bracelet at the end of that day so that you can take it home with you right away. The choice you make is up to you. Remember you are NOT choosing for today so you will get glow bracelets tomorrow or the day after tomorrow.

When you are done, turn your paper over, and wait quietly. I will collect your sheet only after everyone is finished. Okay, GO!

Okay, I will collect the activity sheets now. Remember, your decision was private! Please do not talk about your decision to anyone. Thanks for doing the activity today. Let's go back to class.

B.0.3 Week 1 Day 2

Thanks for doing the activity. First, we are going to sit quietly. It is very important that we do not talk during the activity. If anyone talks, they will have to go back to class and will not get any prizes. Can we practice being quiet? [wait for kids to settle down]. Here are the glow bracelets that we are playing with. When you get the bracelets, they will glow for 1 day. But after 1 day they will stop glowing. (Walk around with tray of glow bracelets).

Yesterday you made a choice about 1 glow bracelet TOMORROW or 2 glow bracelets the DAY AFTER TOMORROW. Since you made that choice yesterday, that means the choices are for 1 glow bracelet TODAY or 2 glow bracelets TOMORROW. Here is the activity sheet you had (Show poster 1). Here were the rules:

- If you said you wanted 1 glow bracelet TOMORROW, that's SOONER, you circled this choice. Now that you slept for one night, that's turned into TODAY (Point to the 1 glow bracelet sheet - use dry eraser maker to cross out TOMORROW and write TODAY, and cross out the sleeping person).

- If you wanted 2 glow bracelets the DAY AFTER TOMORROW, that's LATER, you circled this choice. Now that you slept for one night, that's turned into TOMORROW (Point to the 1 glow bracelet sheet - demonstrate crossing out DAY AFTER TOMORROW and writing TOMORROW and cross out one sleeping person).

If you chose for TODAY (Point to poster 1), you will get your glow bracelet as soon as we finish the activity today. If you chose for TOMORROW (Point to poster 1), you will wait and get your glow bracelets tomorrow before you go home from school.

Okay, I will pass out your activity sheets from yesterday now, and you can see the choice that you made quietly. (Pass out sheets.) Please keep your eyes on your paper only and no talking.

Doing the Activity

(DO NOT COLLECT ACTIVITY SHEETS UNTIL THE END!) Now there is a surprise! You can change your decision if you want. You can either STAY WITH YOUR CHOICE from yesterday or CHANGE YOUR MIND. Let me show you the rules:

There is a new yellow activity sheet that counts as the final decision. You will only get glow bracelets if you follow the rules for the new yellow sheet. (Show yellow poster 2). Here are the rules:

- If you want 1 glow bracelet TODAY - that's SOONER, circle this choice. This shows ONE bracelet. It also shows you will get the bracelet as soon as we finish the activity TODAY. Since, today is TUESDAY/ THURSDAY (Use yellow poster - point to the 1 glow bracelet sheet and each component).

- If you want 2 glow bracelets TOMORROW - that's LATER, circle this choice. This shows TWO bracelets. It also shows that you will have to sleep for ONE night before tomorrow comes. Since tomorrow is WEDNESDAY/ FRIDAY (Point to the 2 glow bracelets sheet and each component)

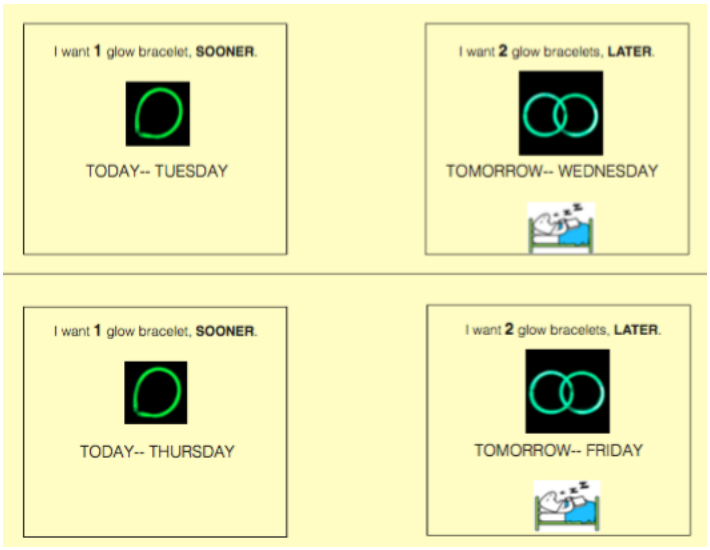


Figure B3: Poster 2 and response sheets: Monday cohort received top panel and Wednesday cohort received bottom panel.

Remember, you are choosing to get 1 glow bracelet TODAY OR 2 glow bracelets TOMORROW - so you should only circle one choice. You will get your glow bracelet in an envelope on the day you decide to get it, and get a blank envelope on the other day. The choice you make is up to you. The choice you make on the new yellow activity sheet is the one that counts as the final decision.

(Pass out activity sheets/pencils. Staple yellow sheet on the child's white sheet.) I will pass out your new yellow activity sheets now. Don't start until I say GO. If you have a question, raise your hand and I will come to you. When you are done, turn your paper over, and wait quietly. I will collect your sheets only after everyone is finished. Okay, GO.

Okay, I will collect the activity sheets now and pass out your envelopes. Remember, your decision was private! Please do not talk about your decision to anyone. (As you collect the activity sheets, check to see what the child chose and give them the corresponding envelope for day 2 - either an envelope with 1 bracelet or an envelope with nothing)

Thanks for doing the activity today. Let's go back to class.

B.0.4 Week 2 Day 1

Thanks for doing the activity. First, we are going to sit quietly. It is very important that we do not talk during the activity. If anyone talks, they will have to go back to class and will not get any prizes. Can we practice being quiet? [wait for kids to settle down]

Here are the glow bracelets that we will play with. (Walk around with tray of glow bracelets).

In the activity, just like last week, you are going to decide WHEN you want to get glow bracelets. When you get the glow bracelets, they will glow for 1 day. But after

1 day they will stop glowing. Today you will decide whether you want 1 glow bracelet TOMORROW - that's sooner, OR 2 glow bracelets the DAY AFTER TOMORROW - that's later. You will get your glow bracelets in an envelope at the end of the day on the day you decide to get it, and get a blank envelope on the other day. The choice you make is up to you.

You will only get glow bracelets if you follow the rules for the activity sheet. Here is the activity sheet you will have (Show poster 1). Here are the rules:

- If you want 1 glow bracelet TOMORROW - that's SOONER, you should circle this choice. This shows ONE glow bracelet. It also shows that you will have to sleep for ONE night before tomorrow comes since tomorrow is TUESDAY/THURSDAY (Point to the 1 glow bracelet sheet as well as each component).
- If you want 2 glow bracelets the DAY AFTER TOMORROW - that's LATER, circle this choice. This shows TWO glow bracelets. It also shows that you will have to sleep for TWO nights before the day after tomorrow comes. Since the day after tomorrow is WEDNESDAY/ FRIDAY (Point to the 2 glow bracelets sheet and each component)

Remember, you are NOT choosing for today and will not get any glow bracelets today.

Doing the Activity I

Okay, I will pass out the activity sheet now. If you have a question, raise your hand and I will come to you. When you are done, turn your paper over, and wait quietly. We will collect your sheets only after everyone is finished. Okay, GO.

Doing the Activity II

(WAIT UNTIL EVERYONE IS DONE, BUT DO NOT COLLECT ACTIVITY SHEETS YET!) Now there is one more thing. Remember, last week you got a surprise: a new yellow activity sheet so you could make your final decision.

Today, you can decide whether you want your final decision to be TODAY - on the white activity sheet that you just filled out OR TOMORROW on the yellow activity sheet. Let me show you the rules. Here is the activity sheet you will have: (Show poster 3).

- If you want your final decision to be on the WHITE sheet that you JUST filled out TODAY, you should circle the white box.
- If you want your final decision to be on the YELLOW sheet that you will fill out TOMORROW, you should circle the yellow box.

Only kids who circle the yellow box will be able to make a new decision tomorrow. Kids who circle the white box will not be able to make a new decision tomorrow, though they will still get a yellow sheet.

Okay, I will pass out the activity sheet now. [pass out sheet.] If you have a question, raise your hand and I will come to you. When you are done, turn your paper over, and wait quietly. We will collect your sheets only after everyone is finished. Okay, GO.

Okay, I will collect the activity sheets now. Remember, your decision was private! Please do not talk about your decision to anyone.

Thanks for doing the activity today. Let's go back to class.

B.0.5 Week 2 Day 2

Thanks for doing the activity. First, we are going to sit quietly. It is very important that we do not talk during the activity. If anyone talks, they will have to go back to class and will not get any prizes. Can we practice being quiet? [wait for kids to settle down]

Here are the glow bracelets that we are playing with. When you get the bracelets, they will glow for 1 day. But after 1 day they will stop glowing. [walk around with tray of glow bracelets].

Yesterday you made a choice about 1 glow bracelet TOMORROW or 2 glow bracelets the DAY AFTER TOMORROW. Since you made that choice yesterday, that means the choices are for 1 glow bracelet TODAY or 2 glow bracelets TOMORROW.

Here is the activity sheet you had [show poster of decision sheet]. Here were the rules:

- If you said you wanted 1 glow bracelet TOMORROW, that's SOONER, you circled this choice. Now that you slept for one night, that's turned into TODAY (point to the 1 glow bracelet sheet -cross out TOMORROW and write TODAY, and cross out the sleeping person).

- If you wanted 2 glow bracelets the DAY AFTER TOMORROW, that's LATER, you circled this choice. Now that you slept for one night, that's turned into TOMORROW (point to the 1 glow bracelet sheet -cross out DAY AFTER TOMORROW and write TOMORROW and cross out one sleeping person).

Remember, you made one more choice. You decided whether to make your final decision on the white activity sheet yesterday or make your final decision on the yellow activity sheet today.

- If you decided make your final decision on the WHITE activity sheet YESTERDAY, you will get a blank yellow sheet and sit quietly.

- If you decided to make your final decision on the YELLOW activity sheet TODAY, you will get a new yellow sheet that will count as the final decision.

Doing the Activity

I will pass out your activity sheets now, and you can see the choices you made quietly. Please keep your eyes on your paper only and no talking. Don't start until I say GO. [pass out sheets and staple yellow w/ child's sticker on top of white sheet].

If you made your final decision yesterday on the white activity sheet, please sit quietly. You will get glow bracelets based on the decision you made yesterday on your white decision sheet.

If you decided to make your final decision TODAY on the YELLOW ACTIVITY SHEET, you will get a new yellow sheet to fill out. You will only get glow bracelets if you follow the rules for the new yellow sheet. Here is the activity sheet you will have [show poster 2]. Here are the rules:

- If you want 1 glow bracelet TODAY - that's SOONER, circle this choice. This shows ONE bracelet. It also shows you will get the bracelet as soon as we finish the activity TODAY. Today is TUESDAY/THURSDAY. (use yellow poster - point to the 1 glow bracelet sheet and each component).

- If you want 2 glow bracelets TOMORROW - that's LATER, circle this choice. This shows TWO bracelets. It also shows that you will have to sleep for ONE night before tomorrow comes. Tomorrow is WEDNESDAY/FRIDAY (point to the 2 glow bracelets sheet and each component)

Remember, you are choosing to get 1 glow bracelet TODAY OR 2 glow bracelets TOMORROW - so you should only circle one choice. You will get your glow bracelet in an envelope on the day you decide to get it, and get a blank envelope on the other day. The choice you make is up to you.

If you have a question, raise your hand and I will answer you privately. When you are done, turn your paper over, and wait quietly. We will collect your sheets only after everyone is finished. Okay, GO.

Okay, I will collect the activity sheets now. Remember, your decision was private! Please do not talk about your decision to anyone. (collect activity sheets and distribute envelopes)

Thanks for doing the activity today. Let's go back to class.